

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE - UFCG
CENTRO DE TECNOLOGIA E RECURSOS NATURAIS-CTR
UNIDADE ACADÊMICA DE CIÊNCIAS ATMOSFÉRICAS - UACA
PROGRAMA DE PÓS-GRADUAÇÃO EM METEOROLOGIA

VALNELI DA SILVA MELO

**ANÁLISE DE FREQUÊNCIA DE SECAS UTILIZANDO ANÁLISE DE
AGRUPAMENTO E DISTRIBUIÇÕES DE PROBABILIDADES**

CAMPINA GRANDE, PB.

JUNHO DE 2016

VALNELI DA SILVA MELO

**ANÁLISE DE FREQUÊNCIA DE SECAS UTILIZANDO ANÁLISE DE
AGRUPAMENTO E DISTRIBUIÇÃO DE PROBABILIDADE**

Dissertação apresentada ao Programa de Pós-Graduação em Meteorologia da Universidade Federal de Campina Grande como requisito para obtenção do título de Mestre em Meteorologia.

Linha de Pesquisa: Meso e Grande Escalas

Orientador: Prof. Dr. FRANCISCO DE ASSIS
SALVIANO DE SOUSA

CAMPINA GRANDE, PB.

JUNHO DE 2016

Ficha catalográfica: elaborada pela biblioteca da ECI

Será impressa no verso da folha de rosto e não deverá ser contada.



Universidade Federal de Campina Grande
Centro de Tecnologia e Recursos Naturais – CTRN
Unidade Acadêmica de Ciências Atmosférica - UACA
Programa de Pós-Graduação em Meteorologia

Dissertação intitulada “**ANÁLISE DE FREQUÊNCIA DE SECAS UTILIZANDO ANÁLISE DE AGRUPAMENTO E DISTRIBUIÇÃO DE PROBABILIDADE**”, de autoria de Valneli da Silva Melo, aprovada pela banca examinadora constituída pelos seguintes professores:

Prof. Dr. FRANCISCO DE ASSIS SALVIANO DE SOUSA - Orientador
UFCG

Prof. Dr. RICARDO ALVES DE OLINDA - Coorientador
UEPB

Prof. Dr. BERNARDP BARBOSA DA SILVA
UFCG

CAMPINA GRANDE, PB.
JUNHO DE 2016

DEDICATÓRIA

Por todo amor, trabalho e carinho, aos meus amados pais,

Walmir Marcolino *in memoriam* e Marlene Rita.

AGRADECIMENTOS

A Deus, por me dar saúde, esperança, alegria e humildade, qualidades indispensáveis para conseguir caminhar na direção certa.

A toda a minha família, que sempre esteve presente nesta caminhada, em particular ao meu esposo Wellington, que sempre me deu força para continuar, e aos meus filhos Gabriel, Luísa e Lívia, por estarem sempre por perto.

Ao Prof. Dr. Francisco de Assis Salviano de Sousa, pela excelente orientação, sempre paciente e encorajador diante das dificuldades encontradas no desenvolvimento desta pesquisa. Agradeço-lhe por todo o tempo disponibilizado, bem como pelos conhecimentos a mim proporcionados, e parableno-o pelo comprometimento, competência e entusiasmo dedicados a este trabalho.

Aos professores do Departamento de Ciências Atmosféricas da UFCG, pela contribuição na obtenção e aprimoramento dos conhecimentos necessários para o desenvolvimento deste trabalho.

Aos novos e velhos amigos, excelentes companheiros de turma, que direta ou indiretamente contribuíram em minha formação acadêmica. Os levarei para sempre em minha vida. E em especial a Adriana, Milena, Thaís, Alécio, Ricardo e Mainar por sempre estarem por perto me ajudando na minha pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pelo apoio financeiro durante a realização desta pesquisa.

A todos, meu muito obrigado.

RESUMO

O objetivo deste trabalho foi o de ajustar funções densidades de probabilidades aos dados das variáveis Severidade e Duração de secas em três sub-regiões do Semiárido do Brasil (SAB). Para tanto, foram utilizados dados de totais mensais precipitados de 320 postos pluviométricos, no período de 1984 a 2014, gentilmente cedidos pela Agência Nacional de Águas (ANA). Foi utilizado o método “RUN” para se obter os dados de Severidade e Duração de secas, a partir dos totais mensais precipitados. Em seguida utilizou-se a técnica de Análise de Agrupamento para regionalizar a Severidade e Duração de secas. O passo seguinte foi obter os ajustes à função distribuição de probabilidades para cada sub-região. As funções distribuição de probabilidades: Gama, GEV e Logística foram as que melhor se ajustaram à variável Severidade de secas, já a os modelos GEV, Weibull com três parâmetros e Gama com dois parâmetros se ajustaram melhor aos dados de Duração de secas para as sub-regiões homogêneas 1, 2 e 3 respectivamente. Os ajustes das variáveis Severidade e Duração de secas à distribuição de probabilidades Normal bivariada não logrou êxito em nenhuma sub-região do SAB.

Palavras-chave: Secas, Análise de agrupamento, distribuição de probabilidade.

ABSTRACT

The objective of this study was to adjust probability density functions to data of severity and duration of droughts in three sub-regions of the semiarid region of Brazil (SAB). To accomplish this task, the totals monthly precipitated of 320 rain gauges were used in the period 1984-2014. These data were kindly provided by the National Water Agency (ANA). Then, was used the method "RUN" to get the severity and duration of droughts based on totals monthly precipitated. Then we used the Cluster Analysis technique to regionalize the severity and duration of droughts. The next step was to get the adjustment of data to Probability Distribution Function (PDF) for each sub-region. The GEV and Logistics were the PDFs that best fitted to the severity of droughts. On the other hands, the GEV models, Weibull three parameters and Gamma with two parameters were adjusted better to the duration of droughts in the sub homogeneous-regions 1, 2 and 3 respectively. Lastly, the severity and duration of droughts were adjusted to normal bivariate distribution. This adjustment was not successful in any subregion of the SAB studied.

LISTA DE FIGURAS

Figura 2.1: Distribuição normal com o SPI tendo uma média de 0 e uma variância de 1.....	32
Figura 3.1: Região Semiárida do Brasil. Fonte: Agência Nacional das Águas (ANA).....	36
Figura 3.2: Fluxograma do SCQ, desde a extração de uma série do banco de dados, a aplicação dos testes, até a verificação ou correção dos dados duvidosos	40
Figura 3.3: Variáveis utilizadas para qualificar a seca. Adaptado de (Dracup et al 1980).	41
Figura 3.4: Representação da normal bivariada	57
Figura 4.1: Distribuição espacial dos postos pluviométricos na área de estudo.....	60
Figura 4.2: Delimitação do Semiárido brasileiro (SAB) em três sub-regiões	62
Figura 4.3: Precipitação média mensal da sub-região 1	63
Figura 4.4: Precipitação média mensal da sub-região 2	64
Figura 4.5: Precipitação média mensal da sub-região 3	65
Figura 4.6: Sub-regiões homogêneas com base nos índices de Severidade de Seca (mm).....	66
Figura 4.7: Variabilidade espacial da Severidade de seca no SAB	68
Figura 4.8: Variabilidade espacial da Duração de secas em meses no SAB	69
Figura 4.9: Função distribuição de probabilidades acumuladas Gama(2) para a sub-região 1.....	72
Figura 4.10: Função distribuição de probabilidades acumuladas GEV para a sub-região 2.....	73
Figura 4.11: Função distribuição de probabilidades acumuladas Logística para a sub-região 3.....	74
Figura 4.12: Função distribuição de probabilidades acumuladas GEV para a sub-região 1.....	76
Figura 4.13: Função distribuição de probabilidades acumuladas Weibull (3) para a sub-região 2.....	77
Figura 4.14: Função distribuição de probabilidades acumuladas Gama (2) para a sub-região 3.....	78
Figura 4.15: Distribuição de probabilidades empírica e teórica da Severidade de seca da sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C)	79
Figura 4.16: Distribuição de probabilidades empírica e teórica da Duração de seca da sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C).....	81
Figura 4.17: Ciclo da variabilidade e da duração e severidade da sub-região 1.....	82
Figura 4.18: Ciclo da variabilidade e da duração e severidade da sub-região 2.....	83

Figura 4.19: Ciclo da variabilidade e da duração e severidade da sub-região 3.....	84
Figura 4.20: Ajuste da distribuição Normal bivariada: (A) sub-região 1 (B) sub-região 2 e (C) sub-região 3.....	86
Figura 4.21: Q-Q Plot bivariado para a sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C)	87

LISTA DE TABELAS

Tabela 2.1 - Classificação do índice de severidade de seca de Palmer (PDSI).	26
Tabela 2.2 - Classificação do índice de anomalia de precipitação.	28
Tabela 2.3 - Classificação do índice de suprimento de água na superfície.	31
Tabela 2.4 - Classificação do índice de precipitação padronizada.	32
Tabela 2.5 - Classificação do Índice de Porcentagem Normal.	34
Tabela 2.6 - Classificação do Índice de Seca de Bhalme & Mooley.	35
Tabela 4.1 - Estatística descritiva dos dados de precipitação do Semiárido brasileiro no período de 1984 à 2014.	61
Tabela 4.2 - Estatísticas descritivas das sub-regiões para a Severidade de seca (mm).	67
Tabela 4.3 - Estatísticas descritivas das sub-regiões para a Duração de secas (meses)	70
Tabela 4.4 - P-valores dos ajustes da Severidade de secas às funções distribuição de probabilidades	71
Tabela 4.5 - P-valores dos ajustes da Duração de secas às funções distribuição de probabilidades	75
Tabela 4.6 - Correlação entre os valores de Duração (meses) e Severidade (mm) e de secas para as três sub-regiões	81
Tabela 4.7 - Teste MVN de Henze-Zirkler.	86

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 Justificativa.....	14
1.2 Objetivos.....	15
1.2.1 Objetivo geral	16
1.2.2 Objetivos específicos.....	16
2 REFERENCIAL TEÓRICO.....	17
2.1 Sistemas precipitantes no Brasil.....	17
2.2 El Nino/Oscilação Sul (Enos).....	17
2.3 Secas no nordeste do Brasil.....	19
2.4 Secas no mundo.....	20
2.5 Tipos de secas.....	22
Seca meteorológica.....	23
Seca agrícola.....	24
Seca hidrológica	24
Seca socioeconômica.....	24
2.6 Índices de secas	25
Índice de severidade de seca de Palmer.....	25
Índice de anomalia de chuvar (IAC).....	27
Índice de umidade da cultura (PALMER, 1968).....	29
Índice de oferta de água à superfície (SHAFER & DEZMEN, 1982).....	29
Índice padronizado de precipitação (SPI).....	31

Índice de porcentagem normal (Percent of Normal – PN)	33
Índice de seca de Bhalme & Mooley (Bhalme & Mooley Drought Index BMDI)	34
3 MATERIAL E MÉTODOS	35
3.1 Região e dados	35
4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS	35
5 CONCLUSÕES	88
6 SUGESTÕES PARA TRABALHOS FUTUROS.....	89
7 REFERÊNCIAS	90
APÊNDICE A – A MANEIRA CERTA DE FAZER.....	98
ANEXO A – COMO APRESENTAR.....	98

1 INTRODUÇÃO

Na maior parte da região Nordeste a precipitação é escassa e tem flutuações interanuais bastante acentuadas, tanto na escala espacial quanto na temporal. Essas flutuações estão intimamente relacionadas com as mudanças nas configurações da circulação atmosférica de grande escala e com a interação oceano-atmosfera no Pacífico e no Atlântico. O impacto causado pelo fenômeno El Niño-Oscilação Sul (ENOS), um exemplo de perturbação climática de escala global, pode ser sentido principalmente pela modificação no regime e no total de precipitação que, dependendo da intensidade do evento, pode resultar em secas severas, interferindo, de forma expressiva, nas atividades humanas. O déficit hídrico é o grande problema enfrentado na região, causando problemas econômicos com graves reflexos sociais.

O semiárido brasileiro que é o alvo principal desse estudo, apresenta média anual de 800 mm. Segundo Melo (2008), por se tratar de uma região com características naturais complexas e altamente heterogêneas em relação à chuva, solo e vegetação, a região semiárida representa um enorme desafio para o uso e o manejo do solo e dos recursos hídricos em sistemas agrícolas sustentáveis.

Nas últimas décadas, a seca tem, cada vez mais, demandado a atenção de ambientalistas, ecologistas, hidrólogos, meteorologistas e agrônomos. De maneira geral, a seca é um fenômeno natural caracterizado pela deficiência de precipitação durante um período prolongado de tempo, que podem ser meses ou anos, resultando na escassez de água para as atividades humanas. Este fenômeno natural e recorrente é considerado um “desastre natural” sempre que ocorre de forma intensa em locais densamente habitados, resultando em danos socioambientais. Dado o crescimento da demanda mundial por água, devido ao crescimento da população, expansão dos setores agropecuários, de energia e de indústrias, esta é uma situação cada vez mais frequente. Desta maneira, os impactos de secas severas causam as maiores perdas econômicas e sociais, incluindo o maior número de pessoas afetadas.

A seca no Nordeste é uma problemática antiga que sempre causou e ainda causa inúmeros transtornos à população, principalmente as de menor poder aquisitivo. O Semiárido brasileiro tem sido caracterizado, desde o início de sua história, pelo estigma da seca.

O primeiro relato de seca no Nordeste do Brasil (NEB) foi em 1583, feito pelo Padre Fernão Cardin, embora esse fenômeno seja tão antigo quanto o surgimento do homem na Terra. A seca (ausência ou deficiência de chuvas) é um fenômeno natural que se diferencia nitidamente de outros fenômenos, tais como: furacões, terremotos, tsunamis e erupções vulcânicas. Enquanto esses últimos iniciam e terminam repentinamente, além de se restringirem,

normalmente, a uma pequena área, a seca, por outro lado, quase sempre inicia lentamente, apresenta longa duração e afeta, na maioria das vezes, uma extensa área (FREITAS, 2005).

A seca é uma característica normal do clima e sua recorrência é inevitável. No entanto, há ainda muita confusão no meio da comunidade científica e na política de tomada de decisão sobre suas características. A seca não tem uma definição que seja aceita em comum acordo por parte de pesquisadores e não há consenso entre a população em geral. A definição de seca varia de acordo com as percepções de cada um e da sua realidade, pois o que a seca significa para um pode não significar a mesma coisa para outra pessoa que tenha uma maneira diferente de olhar.

Desta maneira, os impactos de secas severas causam as maiores perdas econômicas e sociais, incluindo o maior número de pessoas afetadas. Pesquisas têm demonstrado que a falta de uma definição precisa e objetiva, em situações específicas, tem sido um obstáculo à compreensão da seca, que já levou a indecisão e a inação de gestores e tomadores de decisão, dentre outros (WILHITE & GLANTZ, 1985; WILHITE et al., 2007). Em casos graves, a seca pode durar anos, ter efeitos devastadores na agricultura, na pecuária e no abastecimento de água.

A seca, apesar de recorrente e complexa, afeta um grande número de pessoas e as atividades econômicas na região semiárida do NEB é um evento climático que ainda precisa ser mais bem estudado. A base sólida do conhecimento acerca do fenômeno seca é importante para apoiar à tomada de decisões técnicas a fim de atenuar seus efeitos.

Apesar do pressuposto de que a severidade da seca e sua duração sejam independentes, na prática essa suposição apenas simplifica o problema de análise das secas. As características multidimensionais da seca dificultam sua análise univariada, visto que essa forma de análise não revela as relações entre as propriedades da seca (KIM et al., 2003). Por esta razão, as distribuições bivariadas devem ser amplamente utilizadas para caracterizar essas propriedades.

A hipótese desta pesquisa é: “a abordagem dos eventos de seca com base na probabilidade univariada e bivariada das variáveis duração e severidade que fornece estimativas mais confiáveis para representar o fenômeno das secas”.

1.1 Justificativa

A região do Semiárido brasileiro é frequentemente assolada por eventos de secas, afetando um grande número de pessoas e as atividades econômicas na região semiárida do

NEB, por isso esse evento climático ainda precisa ser melhor estudado. A base sólida do conhecimento acerca do fenômeno seca é importante para apoiar à tomada de decisões técnicas a fim de atenuar seus efeitos. Essas são as principais justificativas para propor o estudo das secas na região do Semiárido brasileiro.

Com o desenvolvimento e a finalização deste trabalho espera-se um aprofundamento dos conhecimentos de sobre seca na região estudada. É uma ambição geral que os resultados deste estudo auxiliem na tomada de decisão dos setores de planejamento de recursos hídricos, monitoramento de tempo e clima, agropecuário, turismo, setor social e da defesa civil estabelecendo um entendimento mais amplo na variação espaço-temporal da precipitação e eventos de seca.

1.2 Objetivos

1.2.1 Objetivo geral

O objetivo deste estudo é utilizar uma metodologia baseada em probabilidade univariada e bivariada para estudar e modelar eventos de seca.

1.2.2 Objetivos específicos

- I. Análise de consistência e atualização de séries temporais de 320 postos no período de 1984 à 2014;
- II. Regionalização dos postos pluviométricos, segundo o critério “período chuvoso” e segundo o critério de “anomalias de secas” utilizando Análise de Agrupamento;
- III. Associar as ocorrências de secas, nas sub-regiões obtidas no item II, aos sistemas meteorológicos de grandes e médias escalas e aos efeitos locais.

2 REFERENCIAL TEÓRICO

2.1 Sistemas precipitantes no Brasil

A variabilidade interanual da precipitação na região Nordeste do Brasil (NEB) constitui um dos grandes desafios da meteorologia devido a influência de fenômenos atmosféricos nas mais diversas escalas espaciais e temporais. Embora localizada na região equatorial, mais precisamente entre as latitudes de 1° e 18° 30' S e 34° 20' e 48° 30' de longitude oeste do meridiano de Greenwich (SUDENE, 2012), o NEB possui um clima bastante particular. A distribuição da chuva na região é bastante heterogênea, apresentando uma média pluviométrica abaixo da média pluviométrica equatorial (KOUSKY e CHU, 1978), variando entre 1800 mm na costa leste a 400 mm na região central (SILVA, 2004).

Além da heterogeneidade do regime pluviométrico no NEB, outro fator marcante com relação ao clima desta região é a existência de uma região semiárida (conhecida como Polígono das Secas), que apresenta precipitação média anual abaixo de 600 mm (MOURA e SHUKLA, 1981; SILVA, 2004). Com efeito, NEB apresenta o ciclo anual da precipitação caracterizado pela estação úmida compreendida entre os meses de novembro a abril, e a estação seca compreendendo os meses restantes do ano (MOSCATI et al., 2007). No entanto, o período que apresenta os maiores totais pluviométricos na região do semiárido do NEB é bastante curto, resumindo-se aos meses de março e abril (HASTENRATH, 2012).

Dada à existência de uma região de clima semiárido, a um interesse de pesquisadores, sobretudo dos meteorologistas, no monitoramento climático dessa região, buscando o conhecimento detalhado das condições atmosféricas, bem como dos mecanismos dinâmicos referentes aos fenômenos oceânico-atmosféricos contribuintes para as variações climáticas observadas no NEB.

O NEB é conhecido como uma região com alta variabilidade espacial-temporal da precipitação, esse fato é responsável pela ocorrência de déficit hídrico durante o período de estiagem e condiciona as características climáticas e as variáveis meteorológicas (MOLINON E BERNARDO, 2002).

Molion e Bernardo (2002) apresentaram uma revisão sobre os mecanismos dinâmicos responsáveis pela variabilidade da precipitação sobre o NEB. De acordo com estes autores, a precipitação no NEB é produzida por mecanismos de grande escala, responsáveis por cerca de 30% a 80% da precipitação observada, e mecanismos de meso e microescala. Dentre os mecanismos de grande escala, os autores destacaram a Zona de Convergência Intertropical (ZCIT), principal contribuinte da precipitação sobre as regiões norte e central do NEB,

ocorrendo em termos médios climatológicos de forma mais intensa e regular entre os meses de fevereiro a abril, período no qual a ZCIT situa-se na posição mais ao sul do seu ciclo sazonal; e sistemas frontais atuantes no sul e no leste do Nordeste, ocorrendo com mais frequência no inverno do Hemisfério Sul, além de perturbações de escala sinótica no campo dos ventos Alísios em resposta a entrada de sistemas frontais em latitudes baixas, que através da convergência dos ventos de sul associados aos sistemas frontais, com os ventos de leste, provocam perturbações conhecidas como Distúrbios Ondulatórios de Leste (DOL).

Segundo Silva (2011), a ocorrência desses eventos também está relacionada à confluência dos ventos Alísios de nordeste e sudeste, bem como do aprofundamento para baixos níveis de uma baixa fria ou Vórtice Ciclônico de Altos Níveis (VCAN) na alta troposfera, propagando-se para oeste. Silva (2011) constatou que os DOL apresentam considerável contribuição na precipitação do setor leste do Nordeste, explicando mais de 50% da precipitação observada entre os estados de Alagoas e Rio Grande do Norte durante o período chuvoso da região. Como mecanismos de mesoescala, Molion e Bernardo (2002) destacam os complexos convectivos e as brisas marítima e terrestre, enquanto circulações orográficas e pequenas células convectivas foram destacadas como os fenômenos de microescala. Estes sistemas também foram destacados por Reboita et al. (2010) como os principais sistemas precipitantes na região do NEB ao apresentar uma revisão sobre os regimes de precipitação na América do Sul.

Estudos realizados por Nobre e Molion (1988) indicaram a relação entre a ZCIT e as anomalias de chuva sobre o Norte do Nordeste (NNE). Em anos de seca no NNE, a ZCIT fica bloqueada mais ao norte de sua posição normal. O NNE fica sob uma região de subsidência que inibe a precipitação pluvial. Em anos chuvosos, ao contrário, a ZCIT move-se até cerca de 5° S e torna-se intensa com o aumento da convergência.

Um mecanismo importante de produção de chuva para o sul e leste do Nordeste é a penetração de frentes frias do Hemisfério Sul, entre as latitudes 5° S e 18° S, conforme foi descrito por Kousky (1979). Oliveira (1986), usando imagens de satélites geostacionários, estudou a climatologia entre 1979-1984 e verificou que as frentes frias frequentemente se associam e interagem com a convecção tropical.

De acordo com Varejão e Silva (2013) a atuação dos Vórtices Ciclônicos do Atlântico Norte (VCAN) é um sistema que contribui para o aumento ou redução das chuvas de verão no NEB.

As Ondas de Leste se formam no campo de pressão atmosférica, na faixa tropical do globo terrestre, na área de influência dos ventos alísios e se deslocam da costa da África até o litoral leste do Brasil. Quando as condições oceânicas e atmosféricas estão favoráveis, as Ondas de Leste também provocam chuvas na região semiárida do NEB (FERREIRA & MELLO, 2005).

Em anos de El Niño, quando as águas superficiais da bacia do Pacífico estão mais aquecidas, toda a convecção equatorial também se desloca para o leste, alterando assim o posicionamento da célula de Walker. Devido à circulação atmosférica, o ar quente sobre essa região é empurrado, originando uma célula descendente sobre o Oceano Atlântico, próximo à região Nordeste do Brasil (NEB) e a Amazônia Oriental. Dependendo da intensidade dessa célula e de sua fase de ocorrência, pode haver inibição na formação de nuvens e na descida da ZCIT, com consequência na redução das chuvas na região do NEB (FERREIRA & MELLO, 2005).

2.2 El Niño/Oscilação Sul (ENOS)

Desde o início do século XX, estudos observacionais e de modelagem numérica vêm sendo desenvolvidos no intuito de relacionar anomalias climáticas observadas em regiões remotas do globo com as anomalias climáticas observadas sobre o NEB, sobretudo em relação às anomalias da Temperatura da Superfície do Mar (TSM) e suas implicações na circulação geral da atmosfera. Gilbert Walker foi pioneiro ao publicar estudos relacionados a oscilações de pressão atmosférica ao nível médio do mar entre os oceanos Índico e Pacífico, denominada de Oscilação Sul (WALKER et al,1924). Hoje se sabe que a interação entre a Oscilação Sul e a variação da TSM no Oceano Pacífico Tropical manifesta-se através de um fenômeno acoplado oceano-atmosfera conhecido como ENOS.

A diminuição da intensidade dos ventos equatoriais de leste enfraquece a ressurgência sobre a costa do Peru e, assim, o Pacífico Equatorial leste torna-se anormalmente mais quente. A esta interação oceano-atmosfera foi dado o nome de ENOS, que é composta por uma fase quente (El Niño), caracterizada pelo aquecimento anômalo na região leste do Pacífico Equatorial, e por uma fase fria (La Niña), caracterizada pelo resfriamento anômalo do Pacífico Equatorial Leste.

O fenômeno ENOS tem sido bastante relacionado com as variações climáticas no NEB por meio da circulação anômala de Walker, onde eventos quentes (frios) no Pacífico Equatorial Leste estariam associados a anos anormalmente secos (chuvosos) no NEB. O

entendimento dos padrões atmosféricos que levam às recorrentes situações de seca registradas no NEB é de grande relevância para o desenvolvimento socioeconômico da região, levando-se em consideração a prevenção e a mitigação dos efeitos danosos registrados, devido à qualidade das chuvas em decorrência do fenômeno El Niño. Neste intuito, muitos pesquisadores vêm apresentando a evolução no conhecimento das fontes/sumidouros tropicais de calor, em especial ao fenômeno ENOS, e suas implicações sobre o clima do NEB através de estudos observacionais e de modelagem numérica (KOUSKY et al., 1984; CHAVES e CAVALCANTI, 2001; COELHO e AMBRIZZI, 2002), que confirmam a estreita relação existente entre o fenômeno ENOS e a variabilidade climática no NEB. Ambos ressaltam que anomalias positivas de TSM associadas à fase quente do fenômeno ENOS (El Niño) no Pacífico Tropical estão associadas a anomalias negativas de chuvas na região Nordeste do Brasil.

2.3 Secas no Mundo

De todos os fenômenos naturais durante o século XX, a seca foi o que causou maior impacto negativo (BRUCE, 1994). Secas severas têm sido observadas em todos os continentes, afetando grandes áreas na Europa, na África, na Ásia e na América Latina (Le Comte, 1995).

Nas últimas décadas, a Europa foi afetada por uma série de grandes eventos de seca, principalmente em 1976, 1989, 1991. Mais recentemente, em 2003, ocorreu uma seca prolongada em grande parte da Europa (FEYEN & DANKERS, 2009). O impacto econômico da seca de 2003 foi cerca de 8,7 bilhões de euros (European Communities, 2007).

A região do Sahel, África Ocidental, desde o final da década de 1960 vem passando por uma seca de severidade sem precedentes na história. Por ter impacto devastador sobre essa região ecologicamente vulnerável, esse evento no Sahel deu grande impulso para a criação da Convenção das Nações Unidas de Combate a Desertificação e a Seca (ZENG, 2003).

Com relação ao continente Asiático, a Índia está entre os países mais vulneráveis à seca. Na Índia a seca tem sido relatada, pelo menos uma vez a cada três anos nas últimas seis décadas. Desde meados da década de noventa, secas prolongadas e generalizadas vêm ocorrendo em anos consecutivos e com aumento das frequências (FAO, 2002).

Na China, as secas severas de 1997, 1999 e 2002 causaram grandes perdas econômicas e sociais (ZHANG, 2003). A seca de 1997, por exemplo, resultou em 226 dias sem vazão no rio Amarelo (ZOU et al., 2005).

No Paquistão, as secas são fenômenos recorrentes. A mais grave ocorreu em 1998 e prolongou-se até 2002 e afetou seriamente as produções agrícola e pecuária, com graves consequências para a segurança alimentar e a subsistência de grande parte da população (ASHRAF & ROYTRAY, 2013).

Os impactos das secas nos EUA têm aumentado significativamente e podem ser verificados pela frequência e severidade delas (CHANGNON et al., 2000). Com base nos dados do National Climatic Data Center (NCDC, 2002) cerca de 10% da área total dos EUA foi atingida por secas moderada e severa durante o século XX. O custo médio anual em perdas agrícolas devido à seca nos EUA é estimado entre 6 e 8 bilhões de dólares (AGRAWALA et al., 2001). Em termos econômicos, a seca é o desastre natural mais caro para os EUA (COOK et al., 2007).

Rossi et al. (1992) estudaram os aspectos espaciais da seca, examinando todas suas propriedades. Os autores quantificaram as secas em diferentes locais usando vários tipos de dados hidrológicos (por exemplo, a precipitação, a vazão e os níveis de reservatórios) de cada uma das estações de observação no âmbito da área de estudo. Levando-se em conta a melhor compreensão de padrões espaciais, um modelo de previsão foi desenvolvido e aplicado. Clause & Pearson (1995) apresentaram a relação entre a duração e severidade das maiores secas anuais em vários locais através da análise de regressão linear. Subsequentemente, o estudo utilizou a análise regional de frequência de secas para produzir resultados mais confiáveis para áreas de estudo com dados limitados ou insuficientes.

Estudos de secas com base em análise de frequência bivariada foram realizados para avaliar situações de seca e fornecer critérios para projetos hidrológicos. Yue et al. (1999), Yue (2000) e Yue e Rasmussen (2002) esclarecem os conceitos básicos para a análise de frequência bivariada e ilustram o uso prático das distribuições bivariadas aplicadas à análise de frequência de cheias.

Através de comparação rigorosa com análises de frequências univariadas, Kim et al. (2006) concluíram que a análise de frequência bivariada de Kernel fornece estimativas mais confiáveis para quantis associados à duração e severidade das secas.

2.4 Secas no Nordeste do Brasil

A seca é um fenômeno natural presente em todas as regiões do mundo. De todos os desastres naturais ocorridos, a seca é responsável por 22% dos gastos, 33% do número de pessoas afetadas e 3% do número de mortes. A Organização Meteorológica Mundial (OMS, 1986) define seca como deficiência de chuva. Já a Convenção das Nações Unidas de Combate à Desertificação e à Seca (em inglês UNCCD, 1994) define seca como um fenômeno que ocorre naturalmente, quando a precipitação pluvial for significativamente abaixo do valor médio climatológico esperado.

Uma das mais graves secas que atingiram todo o Nordeste foi a de 1877/1879. O Estado do Ceará, na época, com uma população de 800 mil habitantes foi fortemente atingido. Em consequência 120 mil pessoas migraram para a Amazônia e 68 mil se mudaram para outros Estados. A economia do Estado foi arrasada (CANUTO et al., 2013).

Segundo Villa (2002) a seca de 1877-1879 foi uma das mais terríveis e dizimou cerca de 4% da população nordestina. Castro (2008) afirma que após a seca de 1877-1879 surgiu um mal terrível no Estado do Rio Grande do Norte: aumentou a presença da cobra cascavel (*Crotalos* hórridos). Pessoas, cerca de 500 e animais que escaparam da seca morreram picadas por esse animal.

A seca de 1979/1984 foi uma das mais prolongadas da história do Nordeste, atingiu toda a região. Uma pesquisa da UNESCO apontou que 62% das crianças nordestinas da zona rural, de 0 a 5 anos, viviam em estado de desnutrição aguda (SILVA et al., 2013).

A seca de 1987 atingiu de forma intensa o norte do Nordeste. Neste episódio a perda da produção de grãos no Ceará foi da ordem de 75% (MARENGO et al., 2011). Em 1988, uma seca intensa atingiu o Nordeste. A população faminta promoveu saques a depósitos de alimentos e feiras livres, animais morreram e lavouras foram dizimadas. Com exceção do Maranhão, todos os outros estados do Nordeste foram atingidos, totalizando cerca de cinco milhões de pessoas afetadas (VAREJÃO E SILVA, 2013).

Já a seca de 1997-1998 reduziu em 72% a produção de feijão, milho, arroz, algodão, e mandioca, segundo a Fundação Joaquim Nabuco (FUNDAJ), em uma pesquisa que envolveu 15 municípios de cinco Estados afetados (NAE, 2005). As perdas na pecuária do Nordeste entre 1997 e 1998 foram de 42,2% de bovinos, 37,2% de caprinos, 40,9% de ovinos e 45,7% de suínos (DUARTE, 2002).

Segundo a Companhia Nacional de Abastecimento (CONAB, 2012) a seca registrada no semiárido nordestino em 2012 causou perdas expressivas na pecuária e na agricultura. As

perdas na produção de feijão e o milho no estado da Paraíba foram de 93,5% e 95,7%; no Rio Grande do Norte 89,6% e 90,1% e no Ceará 87,3% e 92,2%, respectivamente.

Ainda não foram publicados dados oficiais sobre a seca de 2012- 2013 no NEB, entre eles a duração, intensidade e gravidade. O que se sabe é que a região foi castigada por severas perdas na agropecuária. Alguns dados preliminares apontam para uma redução no PIB agropecuário da ordem de R\$ 12 bilhões. A falta de chuva que afeta a vida de milhões de brasileiros no Nordeste têm deixado marcas profundas na atividade econômica. E não é só na agricultura e na pecuária, a seca causou indisponibilidade de água em aproximadamente 95% dos reservatórios, o colapso no abastecimento vem provocando uma crise sem precedentes nos municípios.

Muitos estudos realizados nos séculos XVIII e XIX buscaram recomendações para o enfrentamento da seca e de suas consequências, com destaque para a construção de açudes e de estradas; o reflorestamento e a transferência da população que residia no semiárido para outras áreas. No século XX, surge outro discurso, sugerindo a implantação de sistemas de exploração de propriedades agrícolas para assegurar a convivência do homem com a seca, essa convivência é voltada para o uso sustentável dos recursos naturais do semiárido e para a quebra do monopólio de acesso a terra, à água e aos outros meios de produção, entre as ações adotadas para a convivência, podem ser citadas as seguintes: cisternas de placas e calçadão; açudes, barragens subterrâneas, tanques naturais, poços, silagem, fenação, plantação de espécies vegetais resistentes à seca e utilização de raças adaptadas à região.

2.5 Tipos de secas

A seca é percebida de diferentes maneiras em regiões com características climáticas variadas e com diferentes níveis de utilização de água. Isso contribui de forma negativa para uma definição rigorosa de seca que seja aceita universalmente e de um modelo uniforme para abordagem de seu estudo.

Existem diversas maneiras de se definir seca, dependendo principalmente da temática da abordagem. Essas abordagens podem ser relacionadas à precipitação pluvial, à agricultura, à hidrologia e à socioeconômica. Sendo assim, vários autores (WILHITE; GLANTZ,1987; BYUN; WILHITE, 1999; MCKEE et al., 1993) definiram quatro tipos de secas: meteorológica, agrícola, hidrológica e socioeconômica.

Embora as secas estejam classificadas nesses quatro tipos, todas são originadas da deficiência de precipitação, que resulta na falta de água para o desenvolvimento de atividades ou para a sobrevivência dos seres vivos (WILHITE, 2003).

Seca meteorológica

A seca meteorológica é expressa apenas com base no grau de sequeidão (frequentemente em comparação a algum volume “normal” ou médio) e a duração do período seco. É uma medida do desvio da precipitação em relação ao valor normal. Este tipo de seca caracteriza-se pela falta de água induzida pelo desequilíbrio entre a precipitação e a evaporação, a qual depende de outros elementos como a velocidade do vento, insolação, temperatura e umidade do ar.

Definições de seca meteorológicas devem ser consideradas como específicas de uma região, em virtude de que as condições atmosféricas que conduzem à falta de precipitações variam consideravelmente de região para região. (PIRES, 2003).

Seca agrícola

É o tipo de seca associada à falta de água disponível no solo, que seria suficiente para atender às necessidades das plantas. Na agricultura, este tipo de seca refere-se à disponibilidade de água necessária para repor as perdas por evapotranspiração, sendo relacionada com as características das culturas e/ou da vegetação natural, ou seja, dos sistemas agrosilvopastoris em geral (PIRES, 2003).

Seca hidrológica

Nesta categoria, a seca está relacionada com a redução dos níveis médios de água nos reservatórios e com a depleção de água no solo, sendo normalmente defasada da seca meteorológica e agrícola, dado que é necessário um período maior para que as deficiências na precipitação pluvial se manifestem em diversos componentes do sistema hidrológico. A água em sistemas de armazenamento hídrico (por exemplo, reservatórios, rios) é muitas vezes utilizada para objetivos múltiplos e concorrentes, o que complica ainda mais a necessidade de quantificação dos impactos. A concorrência pela água nesses sistemas de armazenamento aumenta durante a seca e os conflitos entre usuários da água crescem significativamente (WILHITE, 1993).

Como as regiões estão interconectadas por sistemas hidrológicos, uma seca à montante pode produzir graves impactos à jusante, à medida que as reservas de água superficial e subterrânea são afetadas. Isto independe se as áreas à jusante estejam sujeitas à seca meteorológica. Mudanças à montante, no uso do solo (por exemplo, desmatamento, alterações

nos padrões de cultivo, minerações) podem alterar o escoamento e as taxas de infiltração no solo, e assim afetar a frequência e a gravidade de secas à jusante (PIRES, 2003).

Seca socioeconômica

Finalmente, a seca socioeconômica associa a oferta e demanda de um bem econômico com elementos da seca meteorológica, hidrológica ou agrícola. Os processos de tempo e espaço referentes à oferta e demanda são os dois processos básicos a serem considerados em uma definição objetiva da seca. Por exemplo, a oferta de um bem econômico (por exemplo, água, alimentação animal, energia hidrelétrica) depende do clima. Na maioria dos casos, a demanda cresce como resultado do aumento da população e/ou do consumo per capita.

Portanto, a seca poderia ser definida como um evento que ocorre quando a demanda supera a oferta, como produto de uma deficiência na oferta de água em função do clima e/ou da super exploração dos recursos hídricos locais. Este conceito de seca reforça a forte simbiose que existe entre a seca e as atividades humanas, enfatizando ainda mais a importância de manejo dos recursos naturais de maneira sustentável (SUASSUNA, 2011).

Resumidamente pode-se dizer que a seca meteorológica pode se desenvolver rapidamente e "acabar abruptamente" enquanto a hidrológica necessita de um período prolongado de déficit de precipitação pluvial, entretanto, persistirá por mais tempo que a meteorológica. Já a seca agrícola pode estar relacionada às fases de desenvolvimento/crescimento dos vegetais. Por fim a seca socioeconômica engloba todos os impactos das anteriores, agindo nos meios de produção de bens e serviços (HEIM JUNIOR, 2002).

2.6 Índices de secas

Os índices de seca assimilam vários anos de variáveis meteorológicas como: precipitação pluvial, temperatura do ar, evapotranspiração, escoamento superficial (runoff), umidade do solo, entre outras variáveis, e assim as combinam a fim de identificar o início de um período de seca. Vários índices de seca foram desenvolvidos para avaliar o efeito e definir diversos parâmetros, como duração, intensidade e severidade (MISHRA & SINGH, 2010). Dentre eles estão o índice de severidade de seca de Palmer (PALMER, 1968); o índice de anomalia de chuva (ROOY, 1965); o índice de umidade da cultura (PALMER, 1968); o índice de oferta de água à superfície (SHAFER & DEZMEN, 1982); o índice padronizado de precipitação (MCKEE et al., 1993) ou *Standardized Precipitation Index – SPI*, em inglês;

Índice de Porcentagem Normal (*Percent of Normal – PN*); Índice de Seca de Bhalme & Mooley (*Bhalme & Mooley Drought Index – BMDI*), dentre outros.

Índice de severidade de seca de Palmer

O Índice de Severidade de Seca de Palmer (PDSI) é um índice de umidade de solo desenvolvido para regiões homogêneas. O índice mede a perda e a demanda de fornecimento de umidade do solo (PALMER, 1968). O objetivo do PDSI é fornecer medidas normalizadas das condições de umidade de modo que comparações usando o índice possam ser feitas entre localidades e entre meses.

Palmer (1968) considera que o total de precipitação pluvial requerido para manter uma área sob condições de economia estável é dependente da média dos elementos meteorológicos e das condições hídricas dos meses precedentes e do atual para a área em questão em um determinado período.

Palmer desenvolveu o PDSI para incluir a duração de uma seca ou um período de umidade. Sua motivação foi a seguinte: um mês anormalmente úmido no meio de um período de seca não tem um impacto maior no índice e uma série de meses com precipitação próxima ao normal seguido a uma seca drástica não significa que a seca acabou. Consequentemente, Palmer desenvolveu critérios para determinar quanto um período de seca ou de umidade começa e termina.

O PDSI responde às condições meteorológicas que têm um estado de anormalidade de seca (ou anormalidade de umidade). O índice é calculado utilizando como dados de entrada todos os termos da equação do balanço hídrico, incluindo evapotranspiração, runoff, recarga do solo e perda de umidade na camada de superfície.

O PDSI geralmente varia de -4,0 a +4,0 e é designado de modo que uma seca extrema (-4,0) em uma divisão climática tenha o mesmo significado em termos de déficit de umidade que uma seca extrema em qualquer outra divisão climática (ALLEY, 1984). A classificação do índice está descrita na Tabela 2.1. O PDSI é calculado tipicamente em base mensal.

Tabela 2.1 - Classificação do índice de severidade de seca de Palmer (PDSI).

PDSI	Classificação
$\geq 4,00$	Extremamente úmido
3,00 a 3,99	Umidade alta
2,00 a 2,99	Umidade moderada

1,00 a 1,99	Umidade baixa
0,50 a 0,99	Umidade inicial
0,49 a - 0,49	Normal
-0,50 a -0,99	Seca inicial
-1,00 a -1,99	Seca suave
-2,00 a -2,99	Seca moderada
-3,00 a - 3,99	Seca severa
$\leq - 4,00$	Seca extrema

As limitações do PDSI envolvem sua inabilidade de caracterizar inteiramente parâmetros hidrológicos, climáticos e geográficos e a variância em cada parâmetro dentro das bacias dos rios (ALLEY, 1984).

O PDSI pode ser obtido pela equação (2.1):

$$PDSI_i = 0,897PDSI_{i-1} + \left(\frac{Z_i}{3}\right) \quad (2.1)$$

sendo:

i = mês de interesse do índice

Z_i = índice de anomalia de umidade.

O índice de anomalia (Z_i) é obtido pela equação (2.2):

$$Z_i = (P_i - \bar{P}_i)K_i \quad (2.2)$$

sendo:

P_i = precipitação observada no mês i ;

\bar{P}_i = precipitação climatologicamente apropriada no mês i , para as condições existentes;

K_i = constante de padronização.

Índice de anomalia de chuvas (IAC)

O Índice de Anomalia de Chuva foi desenvolvido por Rooy (1965) e incorpora um procedimento de classificação para ordenar magnitudes de anomalias de precipitações positivas e negativas. Esse índice é considerado muito simples, dada a sua facilidade de procedimento computacional, que consiste da determinação de anomalias extremas. Rooy (1965) descreve que o índice visa tornar o desvio da precipitação em relação à condição

normal de diversas regiões passíveis de comparação. O procedimento para o cálculo é apresentado nas equações (2.3) e (2.4):

$$RAI = 3 \left[\frac{(p - \bar{p})}{\bar{m} - \bar{p}} \right] \quad (2.3)$$

$$RAI = -3 \left[\frac{(p - \bar{p})}{\bar{x} - \bar{p}} \right] \quad (2.4)$$

sendo:

p = precipitação observada (mm);

\bar{p} = precipitação média (mm);

\bar{m} = representa a média dos dez valores mais altos;

\bar{x} = representa a média dos dez valores mais baixos.

Para anomalias positivas, o parâmetro \bar{m} é a média dos dez valores mais elevados da precipitação do período estudado; para as anomalias negativas, o parâmetro \bar{x} representa os dez valores mais baixos da precipitação do mesmo período.

Os valores do índice são ordenados em um esquema de classificação de nove categorias (Tabela 2.2) variando de extremamente úmido a extremamente seco (ROOY, 1965). Oladipo (1985) encontrou que diferenças entre o Índice de Anomalia de Precipitação e os mais complicados índices de Palmer e Bhalme-Mooley são insignificantes.

Tabela 2.2 - Classificação do índice de anomalia de precipitação.

PDSI	Classificação
$\geq 4,00$	Extremamente úmido
3,00 a 3,99	Umidade alta
2,00 a 2,99	Umidade moderada
0,5 a 1,99	Umidade baixa
0,49 a - 0,49	Normal
-1,99 a 0,5	Seca suave
-2,00 a -2,99	Seca moderada
-3,00 a - 3,99	Seca alta
$\leq - 4,00$	Extremamente seca

Índice de umidade da cultura (PALMER, 1968)

O Índice de Umidade de Cultura (CMI), desenvolvido por PALMER (1968), utiliza uma aproximação meteorológica para monitorar as condições das culturas numa escala semanal. Em comparação com o PDSI, que monitora longos períodos de umidade e seca meteorológica, o CMI foi projetado para avaliar as condições de umidade em um curto período de tempo nas principais regiões de produção agrícola. É baseado na temperatura média e precipitação total semanal dentro de uma divisão climática e incorporar o valor do CMI da semana anterior. O CMI responde rapidamente às condições de mudança e é influenciado pelo local e pelo tempo.

Devido ao CMI ser projetado para monitorar as condições de umidade em curto prazo para o desenvolvimento da cultura, ele não é uma boa ferramenta de monitoramento de seca em longo prazo. Sua resposta rápida para mudanças nas condições em curto prazo pode fornecer informações errôneas sobre condições em longo prazo. Por exemplo, uma precipitação proveitosa durante uma seca pode permitir que o valor de CMI indique condições de umidade adequadas, enquanto a seca em longo prazo persiste no local. Outra característica de limitação é que o CMI tipicamente inicia e termina em cada estação de crescimento próximo a zero. Isso impede que o CMI seja usado para monitorar as condições de umidade fora da estação de crescimento principal, especialmente em secas que se estendem sobre vários anos. Além disso, o CMI pode não ser aplicável durante a fase de germinação da semente.

Índice de oferta de água à superfície (SHAFER & DEZMEN, 1982)

O Índice de Abastecimento de Água na Superfície (SWSI) foi desenvolvido por SHAFER E DEZMAN (1982). Este índice foi formulado para uso em áreas montanhosas onde a neve tem um papel significativo. Os percentis para a quantificação da seca sazonal de precipitação, neve, fluxo em superfície e reservatório de armazenamento são determinados separadamente, combinados em um único índice com pesos e dispostos em uma escala de -4 a +4, típica escala do índice de Palmer.

A questão de como determinar o peso permanece em aberto. Eles precisam variar durante o ano, para estimar os elementos como a neve, que desaparecem no verão, ou para elementos que são pequenos ou valores artificialmente manipulados, como os reservatórios de armazenamento. Como combinar os efeitos de grandes reservatórios com pequena

variabilidade relativa e pequenos reservatórios com grande variabilidade na mesma bacia de drenagem é também um problema. O SWSI é mais sensível a mudanças em valores de classificação próximo ao centro de sua escala e menos sensível próximo aos extremos.

O objetivo do índice SWSI é incorporar características hidrológicas e climatológicas em um único valor de índice, assemelhando-se ao índice de Palmer. Os valores do índice podem ser normalizados para permitir comparações entre bacias. São necessários quatro valores de entrada de dados dentro do SWSI: bloco de neve, fluxo em superfície, precipitação e armazenamento de reservatório. Devido ao fato desse índice ser dependente da estação do ano, no inverno o SWSI é calculado somente com dados de neve, precipitação e armazenamento do reservatório.

Durante os meses de verão, o fluxo de superfície substitui o bloco de neve como um componente dentro da equação do SWSI.

Para determinar o SWSI para uma bacia hidrográfica em particular, dados mensais são coletados e somados para todas as estações de precipitação, reservatórios e fluxo em superfície/bloco de neve em estações de medição sobre a bacia. Cada componente é normalizado usando análise de frequência acumulada de uma série de dados históricos. A probabilidade de não excesso (a probabilidade que somas subsequentes da qual a componente não excederá a soma corrente) é determinada para cada componente com base na análise de frequência. Isso permite que comparações das probabilidades sejam feitas entre as componentes.

Cada componente em um peso é atribuído a ele, dependendo de sua contribuição típica para a água em superfície dentro dessa bacia. Esses componentes pesados são somados para determinar um valor de SWSI representando a bacia inteira. A equação (2.5) representa o índice SWSI.

$$SWSI = \frac{aP_{snow} + bP_{prec} + cP_{strm} + dP_{resv} - 50}{12} \quad (2.5)$$

sendo:

a, b, c, d = pesos para neve (snow), precipitação (prec), fluxo em superfície (strm) e armazenamento de reservatório (resv), respectivamente.

P_i = probabilidade (%) de não excesso para cada um destes quatro componentes do balanço hídrico.

Os cálculos são executados com um escala de tempo mensal. Para cada mês, os valores de cada componente medido em todas as estações (ou reservatórios) através da

região/base são somados. Cada soma é normalizada e sua probabilidade de não excesso é determinada. Os pesos são atribuídos a cada componente do balanço hídrico, dependendo de sua contribuição típica para a água em superfície dentro de uma bacia. Assim como o índice de Palmer, o SWSI é centrado em zero e tem uma escala entre -4 e 4 como podem ser observados na Tabela 2.3.

Tabela 2.3 - Classificação do índice de suprimento de água na superfície.

SWSI	Classificação
- 4,00 a -3,00	Extremamente seco
- 2,99 a -2,00	Seca moderada
-1,99 a -1,00	Seca inicial
-0,99 a 0,99	Normal
1,00 a 1,99	Suprimento acima do normal
2,00 a 2,99	Suprimento moderado
3,00 a 4,00	Suprimento abundante

Índice padronizado de precipitação (SPI)

O déficit de precipitação possui diferentes impactos na água subterrânea, no armazenamento do reservatório, na umidade do solo e nos escoamentos dos rios. Isso fez com que McKee et al. (1993) desenvolvessem o Índice de Precipitação Padronizado (SPI). Esse índice quantifica o déficit de precipitação para múltiplas escalas de tempo que refletem o impacto da seca na disponibilidade de fontes de água. As condições de umidade do solo respondem às anomalias de precipitação em uma escala de tempo relativamente curta. O armazenamento de água subterrânea, dos fluxos de rios e do reservatório reflete as anomalias de precipitação em longo prazo. Por essa razão, MCKEE et al. (1993) originalmente calcularam o índice em escalas de tempo de três, seis, doze, vinte e quatro e quarenta e oito meses.

O cálculo do índice SPI para qualquer local é baseado no registro da precipitação de longo prazo ajustado a uma distribuição de probabilidade. Essa distribuição é então transformada para uma distribuição normal, de modo que o SPI médio para uma localização e período desejados seja zero (EDWARDS; MCKEE, 1997). Os valores positivos de SPI

indicam precipitação maior do que a mediana ou média e os valores negativos indicam precipitações menores que a mediana ou média.

Devido ao fato do SPI ser normalizado, climas mais secos ou úmidos podem ser representados no mesmo modo. Assim, períodos úmidos podem também ser monitorados usando o SPI.

Para o cálculo do SPI, deve-se utilizar uma base de dados de precipitação com pelo menos 30 anos, sendo ajustado por meio da distribuição gama. A distribuição de probabilidade cumulativa $H(x)$ é então transformada em uma variável aleatória normalizada (Z) com média zero e desvio padrão 1, onde a variável (Z) corresponderá ao valor de SPI.

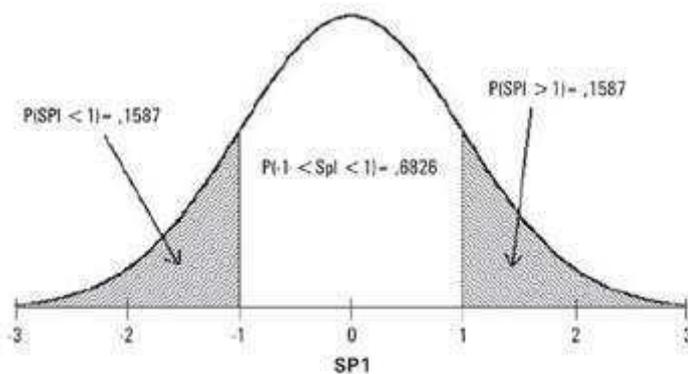


Figura 2.1: Distribuição normal com o SPI tendo uma média de 0 e uma variância de 1.

Segundo McKee et al. (1993), a seca ocorre sempre que o valor do SPI é continuamente negativo, atingindo uma intensidade igual ou menor que menos um (-1,0). A seca termina quando o valor do SPI se torna positivo. A classificação é feita com base nos limites indicados na Tabela 2.4, permitindo caracterizar não somente as secas, mas também os períodos mais úmidos.

Tabela 2.4 - Classificação do índice de precipitação padronizada.

SPI	Classificação
$\geq 2,00$	Extremamente úmido
1,00 a 1,99	Muito úmido
0,50 a 0,99	Moderadamente úmido
0,49 a -0,49	Próximo ao normal
-0,50 a -0,99	Moderadamente seco

-1,00 a -1,99	Muito seco
≤ -2,00	Extremamente seco

Índice de porcentagem normal (Percent of Normal – PN)

O índice de Porcentagem Normal (PN) é facilmente calculado e serve bem às necessidades dos previsores do tempo e do público em geral. Esse índice é baseado na precipitação, que é facilmente medida em diversos locais. Tais análises são eficazes quando usadas para uma única região ou uma única estação. O cálculo é feito dividindo a precipitação atual pela precipitação normal (considerada a média de 30 anos) e multiplicado por 100, para obter o resultado em porcentagem, conforme mostra a equação 1. Esse índice pode ser calculado em diversas escalas de tempo, como semanas, décadas, meses, sazonal e anual. Tem como grandes vantagens a simplicidade e a facilidade de determinação e a existência de dados disponíveis na maioria das estações.

Uma desvantagem potencial de se usar esse índice é quanto à distribuição da precipitação no período considerado, que não obedece a uma distribuição normal. Nessas circunstâncias, há uma diferença entre a média e a mediana, sendo a mediana o valor com uma probabilidade e ocorrência de 50%. Isso ocorre porque a maioria dos valores hidrológicos aleatórios em escalas mensais ou sazonais não apresenta uma distribuição normal.

$$P = \left[\frac{\text{Prp}(atual)}{\text{Prp}(normal)} \right] * 100\% \quad (2.6)$$

sendo:

P= porcentagem da precipitação normal (%);

Prp (atual) = precipitação medida em um dado local ou estação (mm);

Prp (normal) = precipitação normal do local ou estação considerada (mm).

A Tabela 2.5 ilustra a classificação dos valores do índice de Porcentagem Normal divididos em onze categorias, adaptada por GOIS (2005).

Tabela 2.5 - Classificação do Índice de Porcentagem Normal.

PN	Classificação
$\geq 100\%$ a 90%	Extremamente úmido
70% a 89,9%	Umidade alta
50% a 69,9%	Umidade moderada
30% a 49,9%	Umidade baixa
10% a 29,9%	Umidade inicial
$9,9\% \leq 0 \leq 9,9\%$	Normal
-10% a -29,9%	Seca inicial
-30% a -49,9%	Seca suave
-50% a -69,9%	Seca moderada
-70% a -89,9%	Seca severa
-90% a -100%	Seca extrema

Índice de seca de Bhalme & Mooley (Bhalme & Mooley Drought Index – BMDI)

BHALME E MOOLEY (1980), ao estudarem as regiões tropicais da Índia, encontraram alguns problemas para aplicar o Índice de Severidade de Seca de Palmer. Assim como outros estudos (HAVENS, 1969; ALLEY, 1984; GUTTMAN, 1991) demonstraram, o índice de Palmer não era um bom indicador das condições de umidade, principalmente nos períodos secos. Outra desvantagem encontrada no PDSI foi que a regularização da vazão superficial não é considerada.

Com os problemas encontrados, Bhalme e Mooley desenvolveram o Índice de Seca de Bhalme & Mooley (BMDI) para avaliar a intensidade de seca usando somente a precipitação. Os detalhes do cálculo do BMDI e do PDSI são geralmente semelhantes, com apenas algumas diferenças. Bhalme e Mooley substituíram o índice (Z) de umidade no algoritmo de Palmer por um índice mais simples de precipitação mensal (M), calculado somente dos dados de precipitação. Além disso, o BMDI não envolve o seguimento simultâneo de pseudos índices dos períodos de seca, úmidos ou não estabelecidos, que poderiam ser confusos.

Para obter o BMDI, primeiramente calcula-se a precipitação média mensal e o desvio padrão, a seguir, se obtém o índice (M) de umidade.

Tabela 2.6 - Classificação do Índice de Seca de Bhalme & Mooley.

BMDI	Classificação
$\geq 4,00$	Extremamente úmido
3,00 a 3,99	Umidade alta
2,00 a 2,99	Umidade moderada
1,00 a 1,99	Umidade leve
0,99 a -0,99	Próximo ao normal
-1,00 a -1,99	Seca leve
-2,00 a -2,99	Seca moderada
-3,00 a -3,99	Seca severa
$\leq -4,00$	Extremamente seco

3 MATERIAL E MÉTODOS

3.1 Região e dados

A Região Semiárida do Brasil é delimitada por uma área de 895.254,4 km². Essa Região está contida no Polígono das Secas. Fazem parte dessa Região 1.042 municípios, distribuídos nos seguintes estados: Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia e Minas Gerais. Essa Região representa 57,53% da área do Nordeste e compreende uma população de 19.326.007 habitantes (40,54% da população do NE), com 56,52% residindo em áreas urbanas e 43,48% em áreas rurais.

Em termos de condições edafoclimáticas, a Região apresenta as seguintes características: temperatura média anual entre 23 e 27 °C, insolação média anual de 2800 horas, umidade relativa do ar média em torno de 65% e precipitação pluviométrica média anual igual ou inferior a 800 mm. Na Região há vários regimes de chuva, caracterizando a irregularidade espaço-temporal do fenômeno. É frequente a ocorrência de secas, caracterizadas pela ausência ou escassez de chuvas (STRANG, 1979).

O solo é litólico com baixa profundidade e substrato predominantemente cristalino. A capacidade de retenção de água desse solo é baixa, conseqüentemente a maioria dos tipos de rios são efêmeros, ou seja, existem somente quando fortes chuvas acontecem. As águas subterrâneas estão distribuídas em bacias sedimentares ou no cristalino, a vegetação dominante da região é a Caatinga.

A Figura 3.1 mostra a área atualizada da Região semiárida do Brasil segundo o estudo do Grupo de Trabalho Interministerial para Redelimitação do Semiárido nordestino e do Polígono das Secas: Relatório final, janeiro de 2005.



Figura 3.1: Região Semiárida do Brasil. Fonte: Agência Nacional das Águas (ANA)

A metodologia descrita a partir do subcapítulo 3.2. foi aplicada aos valores dos totais mensais precipitados de 320 postos pluviométricos, com tamanho amostral superior a 30 anos, localizados no âmbito da Região Semiárida do Brasil, listados nas Tabelas A.1 em Anexo. Os dados utilizados aqui foram obtidos pela Agência Nacional de Águas (ANA), Instituto Nacional de Meteorologia (INMET) e Agência Executiva de Gestão das Águas do Estado da Paraíba (AESPA).

No entanto, um grande fator de dificuldade é a qualidade das séries de precipitações que quase sempre apresentam falhas que muitas vezes inviabilizam a investigação para determinada localidade, ou para determinados índices.

Para este estudo, os dados passaram por dois procedimentos essenciais: preenchimento de falhas de dados e controle de qualidade. Cada um destes procedimentos será melhor explicado a seguir.

Preenchimento de falhas em dados meteorológicos

Em Meteorologia e Climatologia, diversos produtos operacionais de monitoramento em tempo e clima dependem de uma boa rede de observações em superfície, que não apresentem ou tenham o menor número possível de falhas. Como exemplo, as séries de dados pluviométricos completas servem tanto para caracterizar a climatologia das chuvas de uma região, como também para estudos de mudanças, tendências e extremos climáticos (SALVADOR, 2010). Lucio et al. (2007) observaram que, frequentemente, dados de registros climatológicos de localidades onde não há observações meteorológicas, têm que ser calculados via interpolações espaço-temporal de observações em estações vizinhas.

Para preenchimento de falhas em dados diários de variáveis meteorológicas foi utilizada a técnica conhecida como MICE (Multivariate Imputation by Chained Equations). Esta é uma técnica de atribuição múltipla que tem uma série de vantagens sobre outros métodos para tratar dados faltantes em séries temporais (SCHAFER e GRAHAM, 2002).

Vários são os campos de aplicação da técnica MICE (VAN BUUREN e GROOTHUIS-OUDSHOORN, 2011), principalmente a dados da área de saúde pública. Nesta pesquisa, ela será adaptada para preenchimento de dados faltosos em séries temporais de variáveis meteorológicas.

O MICE é desenvolvido em linguagem R, em forma de pacote, para versões maiores que a 2.10. Essa técnica é preenche dados faltantes em series históricas, faltas essas causadas por motivos diversos, como falha no instrumento ou problemas nas estações meteorológicas. Os dados faltantes são caracterizados por “NA”, e foi usado o número de imputações múltiplas padrão do pacote MICE versão 2.12 (m=5 iterações) do software estatístico livre R, as imputações são geradas de acordo com o método padrão, que é, para dados numéricos. São criados vários conjuntos de dados completos, baseados em regressões lineares, regressões logísticas, modelos multinomiais log-lineares, ou regressões de Poisson para cada variável, como seja requerido, e os dados ausentes são imputados com base nos valores observados e suas relações com os dados de cada conjunto, envolvendo a criação de várias previsões para cada valor em falta, sendo, finalmente, usados os dados do conjunto que produziu menos incertezas e menores erros, entre estes e a série de dados observada (GREENLAND e FINKLE, 1995).

Sistema de controle de qualidade em dados meteorológicos

Todos os sistemas de monitoramento do clima e conjuntos de dados climáticos exigem a melhoria da qualidade dos dados, continuidade e homogeneidade. Isso vale para praticamente todas as aplicações climatológicas, mas torna-se uma necessidade essencial em termos da capacidade de medir, detectar ou atribuir mudanças climáticas. Para dados diários, onde a acessibilidade ainda é muito restrita, Sistemas de Controle de Qualidade (SCQ), particularmente rigorosos, são essenciais (WMO, 2003; CAMARGO e HUBBARD, 1999). Os dados não devem ser considerados como satisfatórios para arquivamento permanente até que tenham sido submetidos a um nível adequado de controle de qualidade.

Um bom SCQ deve combinar dois fatores: pessoas e máquinas. Um software de qualidade de dados deve fornecer uma lista de dados suspeitos, mas as decisões finais sobre a correção ou atualização do arquivo digital devem ser do pessoal competente que estará analisando tais relatórios. No entanto, a massa de dados disponíveis nos Serviços Meteorológicos Nacionais (particularmente no INMET) só aumenta, por exemplo, com a inserção de dados provenientes de estações meteorológicas automáticas de alta resolução. No entanto, a contrapartida geralmente está associada a restrições de pessoal e na qualidade da verificação automática.

Um SCQ não deve ser uma ferramenta perfeita, que acusa um possível erro e automaticamente o dado é retirado da série. Algumas técnicas de SCQ, com forte componente estocástica, podem levar a uma alta probabilidade de rejeição de boas observações (GUTTMAN, 1990; MEEK, 1994). Um SCQ deve disponibilizar ao tomador de decisão, saídas gráficas facilmente compreensíveis e resumos contendo a lista de dados suspeitos, que são excelentes ferramentas para exames visuais que podem ser de grande ajuda, uma vez que a mente humana é especialmente hábil em identificar padrões espaciais.

Alguns testes de consistência são um importante conjunto de verificações de possíveis erros, pois devem explorar a inter-relação temporal e espacial dos dados climatológicos. Os três principais meios de verificações de consistência são os internos, os temporais, e os espaciais:

- Os testes de consistência interna exprimem as relações físicas entre os diferentes elementos climatológicos. Em alguns casos são testes lógicos do tipo: se determinado elemento encontra-se em algum intervalo, outro deverá estar contido, também, em outro determinado intervalo (GANDIN, 1988);

- testes de consistência temporal são baseados na persistência no tempo de elementos climatológicos. Certos limiares de mudanças selecionadas dependem da variável em questão, do período do ano e da região climática a que pertence os elementos da série temporal (EISCHEID et al., 1995);
- testes de consistência espacial exploram a suave variação no espaço das variáveis climatológicas. Geralmente este tipo de teste envolve a estimativa de um determinado elemento com base nas observações vizinhas em uma mesma região climática (HUBBARD et al., 2005). O limite de diferenças aceitas dependerá do tipo de variável, da região climática e da distancia entre as estações. Logo, a eficácia deste tipo de teste vai depender da disponibilidade de estações vizinhas (YOU et al., 2007).

O SCQ aplicado neste estudo é baseado em uma série de testes, denominados Grupos de Testes. O fluxograma apresentado na Figura 3.2, detalha o passo a passo de cada etapa à qual todos os registros de dados são analisados. Todas as colunas com informações são cuidadosamente estudadas, desde a coluna que contém o código de identificação da OMM da estação, a coluna que contém as datas da coleta dos dados, e os respectivos valores das variáveis meteorológicas. Uma etapa importante é a confecção de um arquivo chamado metadados, que contém as informações básicas da estação: o código de identificação da OMM, o nome da estação, a longitude e a latitude em graus, minutos e segundos, assim como em décimos de graus, a altitude, o País e Estado a qual pertence, a data de início e fim de suas operações, a Instituição a qual pertence e o tipo da estação, se convencional ou automática. No pré-processamento do SCQ, estas informações contidas nos metadados são lidas e servem de base para alguns dos testes gerais, apresentados na Figura 3.2:

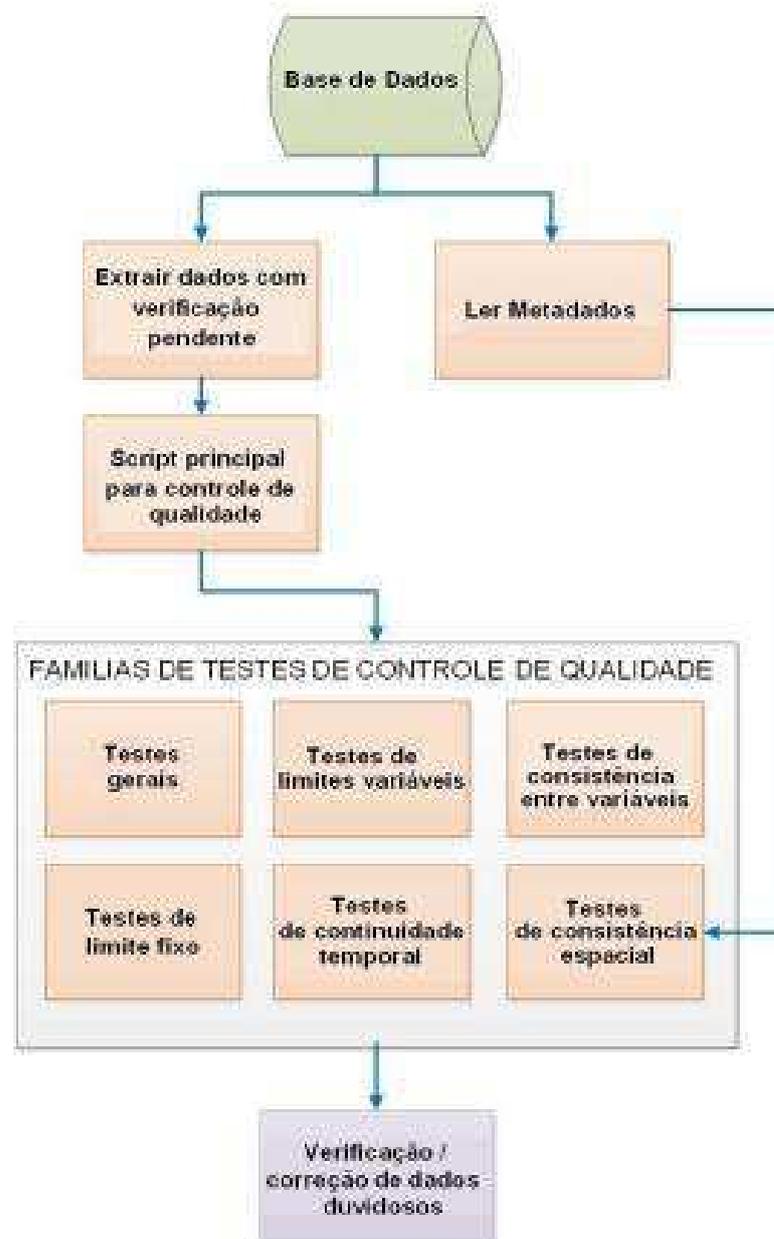


Figura 2.2: Fluxograma do SCQ, desde a extração de uma série do banco de dados, a aplicação dos testes, até a verificação ou correção dos dados duvidosos

3.2 Definições dos parâmetros de seca

Um índice de seca é uma variável que avalia o seu efeito sobre determinada atividade e define diversos parâmetros da seca. A Figura 3.3, ilustra a representação detalhada das

propriedades de seca, como a duração, intensidade e severidade. Neste estudo, a duração da seca e severidade foi extraída de séries temporais de precipitação mensal para cada estação.

A duração é o período de tempo entre o início e o fim de uma seca (período de tempo em que sucessivas chuvas são menores do que a média climatológica), pode ser expressa em semanas, meses ou anos.

Já a severidade de seca é a soma do déficit (período de tempo em que sucessivas chuvas são menores do que a média climatológica) para a duração de seca. A intensidade de seca é dada pela razão entre a severidade e a duração. Um índice de seca deve ser capaz de identificar uma seca em diferentes escalas de tempo.

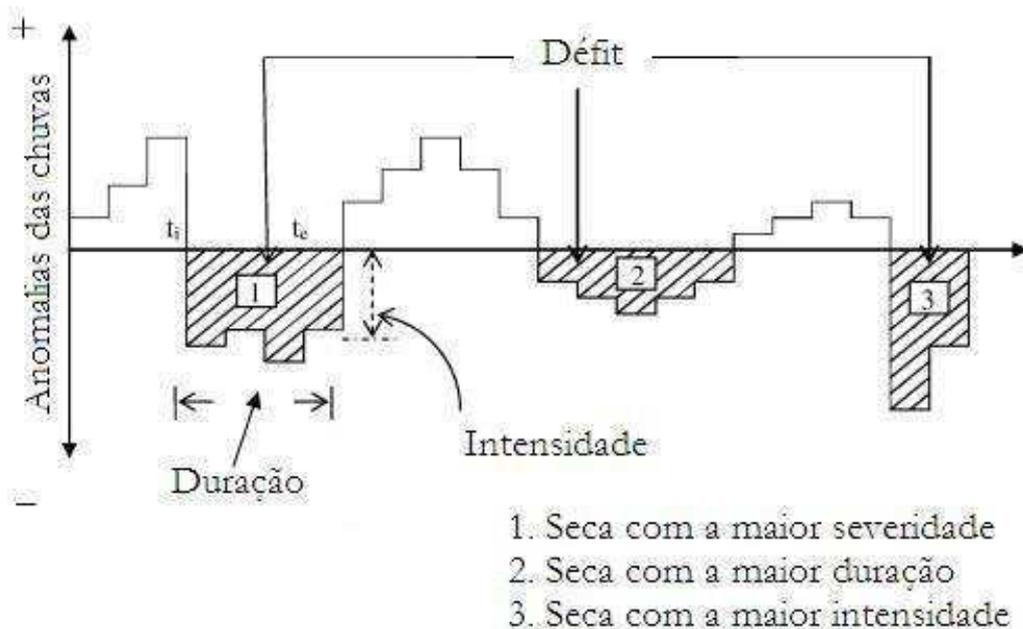


Figura 3.3: Variáveis utilizadas para qualificar a seca. Adaptado de (Dracup et al 1980).

Neste estudo, a duração e a severidade de secas foram extraídas das anomalias das séries temporais de totais mensais precipitação de cada posto pluviométrico. Foram também calculadas várias estatísticas descritivas básicas para as propriedades duração e severidade (por exemplo, média, valor máximo, valor mínimo, variância, coeficiente de assimetria e o número de eventos de seca). Ajustou-se uma função distribuição de probabilidade bivariada para as propriedades duração e severidade para cada local. Se os coeficientes de correlações entre a duração e a severidade da seca forem significativos na maioria dos locais poderá se optar pela análise bivariada regional.

As incertezas na análise de frequência pontual de seca ainda são problemáticas porque secas extremas são raras, por definição, conseqüentemente, suas amostras, muitas vezes, são pequenas para estimativas confiáveis de valores extremos. O principal objetivo deste estudo é estabelecer uma forma de superar as limitações da análise de frequência pontual de seca, aumentando o tamanho amostral dos dados na área de estudo. Uma forma importante para aumentar o tamanho das amostras é agrupar postos pluviométricos que têm propriedades semelhantes de secas. Essa abordagem melhora a precisão da análise de seca, por meio da análise de frequência regional, haja vista que reduz os erros de amostragem das propriedades de seca.

3.3 Delimitações das sub-regiões por análise de agrupamento

A análise de agrupamento (AA) é uma ferramenta útil para a análise de dados em muitas situações diferentes. Esta técnica pode ser usada para reduzir a dimensão de um conjunto de dados, reduzindo uma ampla gama de objetos à informação do centro do seu conjunto.

A análise de agrupamento pode ser usada não só para identificar uma estrutura presente nos dados, mas também para impor uma estrutura num conjunto de dados mais ou menos homogêneos que têm de ser separados.

Em especial, a AA, utilizada em processos de classificação, consiste em determinar o nível de similaridade ou dissimilaridade entre indivíduos, aplicando uma função de agrupamento a uma determinada variável, transformando um conjunto heterogêneo de unidades, neste caso estações meteorológicas, em grupos caracterizados pela homogeneidade interna e encontrando-se, também, grande número de pela heterogeneidade externa (JOHNSON, WICHERN, 2007).

A distância euclidiana é a medida de distância mais frequentemente empregada quando todas as variáveis são quantitativas. A distância euclidiana é utilizada para calcular medidas específicas, assim como a distância euclidiana simples e a distância euclidiana quadrática ou absoluta, que consiste na soma dos quadrados das diferenças, sem calcular a raiz quadrada. A distância euclidiana quadrática é definida por:

$$D_e = \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \quad (3.1)$$

onde:

x_{ij} é a j -ésima característica do i -ésimo indivíduo;

$x_{i'j}$ é a j -ésima característica do i' -ésimo indivíduo.

Quanto mais próximo de zero for a distância euclidiana, mais similares são os objetos comparados.

Procedimentos de aglomeração

A aglomeração hierárquica se caracteriza pelo estabelecimento de uma hierarquia ou estrutura em forma de árvore. A aglomeração hierárquica interliga os objetos por suas associações, produzindo uma representação gráfica chamada de dendrograma, onde os objetos semelhantes, segundo as variáveis estudadas, são agrupados entre si. Já na aglomeração não hierárquica, assume-se um centro de agrupamento e, em seguida, agrupam-se todos os objetos que estão a menos de um valor pré-estabelecido do centro.

Para os procedimentos de aglomeração deve-se optar por um método específico. Neste trabalho serão utilizados o método Ward para aglomeração hierárquica.

Método Ward

É um método de partição que fornece indicações mais precisas sobre o número de conglomerados a ser formado. Este método talvez seja um dos mais utilizados quando se têm muitos objetos para agrupar, com pequenas variações. O critério mais utilizado de homogeneidade dentro do grupo e heterogeneidade entre os grupos é o da soma dos quadrados residual baseado na Análise de Variância. Assim, quanto menor for este valor, mais homogêneos são os elementos dentro de cada grupo e melhor será a partição (BUSSAB et al., 1990).

O método de Ward forma grupos, minimizando a dissimilaridade, ou minimizando o total das somas de quadrados dentro de grupos, também conhecida como soma de quadrados dos desvios (SQD).

Em cada etapa do procedimento, são formados grupos, de tal maneira que a solução resultante tenha a menor SQD dentro de grupos. Nessas etapas, são consideradas as uniões de todos os possíveis pares de grupos e, os dois que resultam em menor aumento de SQD são agrupados até que todos os grupos formem um único, reunindo todos os indivíduos (EVERITT & DUNN, 1991).

Valores médios da precipitação mensal dos municípios foram organizados na forma de matriz $P_n \times p$, em que o elemento P_{ij} representava o valor da i -ésima variável (localidade) do

j -ésimo indivíduo (mês). Dessa forma, cada vetor linha representava uma localidade e cada vetor coluna um mês.

Depois do agrupamento dos posto pluviométricos com precipitação similar, aplicou-se novamente o método de Ward para determinar, dentro de cada um dos períodos mensais definidos, as regiões homogêneas de localidades com similaridade da precipitação. Nesse caso, a matriz $P_n \times p$ foi organizada de forma que os indivíduos representavam as localidades e os meses agrupados (períodos), as variáveis.

Teste de heterogeneidade das sub-regiões

HOSKING & WALLIS (1997) apresenta uma formulação para medir heterogeneidade (H). O método compara a dispersão do L-momentos da amostra de indivíduos com os L-momentos de cada um dos possíveis grupos, de modo que H é usada para avaliar se as sub-regiões podem ser razoavelmente tratadas como sub-regiões homogêneas.

A estatísticas para avaliar a heterogeneidade (H) é dada por:

$$H = \frac{V - \mu_v}{\sigma_v} \quad (3.2)$$

e

$$V = \sqrt{\left\{ \frac{\sum_{i=1}^N n_i (\tau^i - \tau^R)}{\sum_{i=1}^N n_i} \right\}} \quad (3.3)$$

Em que μ_v , σ_v é a média e o desvio-padrão dos valores simulados de V, respectivamente, n_i é o tamanho da amostra do local i , τ^i τ^R é o L-momento amostral do local i , e τ^R é a média regional amostral do L-momentos. O L-coeficiente de variação, o L-coeficiente de assimetria, e o L-coeficiente de curtose são utilizados como os L-momentos na Eq.(3.3) para calcular H(1), H(2) e H(3), respectivamente. Uma sub-região é considerada aceitavelmente homogênea quando $H < 1,0$ possivelmente heterogêneas se $1,0 < H < 2,0$, e definitivamente heterogênea se $H > 2,0$.

3.4 Distribuições de probabilidade

Um conjunto de dados climatológicos necessita ser previamente analisado com base em alguns indicadores estatísticos básicos para que se possa, efetivamente, desenvolver a teoria das probabilidades às situações práticas desejadas. Primeiramente, este conjunto de dados é conhecido como série histórica e consiste, basicamente, de uma amostra extraída de uma população.

Com base nesta amostra, pode-se calcular alguns indicadores e medidas estatísticas importantes, como média, desvio padrão (variância), assimetria, curtose e distribuição de frequência dos dados observados na amostra. Inicialmente busca-se a função de densidade de probabilidade e na sequencia usa-se a Função de Distribuição.

Distribuição de probabilidade Qui-quadrado

Quando, na distribuição gama, o parâmetro de forma é igual a $\frac{n}{2}$, com n inteiro positivo, e o parâmetro de escala é $\beta = 2$ resulta a distribuição qui-quadrado com n graus de liberdade cuja densidade é dada por:

$$f_x(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad \text{se } x > 0 \quad (3.6)$$

Usa-se então a seguinte notação para indicar que X tem distribuição qui-quadrado com n graus de liberdade $X \sim \chi_n^2$:

$$X \sim \chi_n^2 \Rightarrow \begin{cases} E(X) = \frac{n}{2} \times 2 = n \\ \text{Var}(X) = \frac{n}{2} \times 2^2 = 2n \end{cases} \quad (3.7)$$

Distribuição de probabilidade Erlang

Quando o parâmetro de forma α da densidade gama é um inteiro positivo, a distribuição n gama é conhecida como distribuição de Erlang e sua função densidade é do tipo:

$$f_x(x) = \begin{cases} \frac{x^{k-1} e^{-\frac{x}{\beta}}}{(k+1)!}, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases} \quad (3.8)$$

Distribuição de probabilidade 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).

$$f_x(x) = \begin{cases} \exp\left(-\frac{x}{\bar{x}}\right), & \text{para } x > 0 \\ 0, & \text{para } x < 0 \end{cases} \quad (3.9)$$

sua função de distribuição acumulada é do tipo:

$$F_x(x) = \int_0^{\infty} f(x) = 1 - \exp\left(-\frac{x}{\bar{x}}\right) \quad (3.10)$$

O único parâmetro da distribuição é a média amostral.

Distribuição de probabilidade gama

Se X for uma variável aleatória contínua e independente, 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:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad \text{se } 0 < x < \infty \quad (3.11)$$

Sendo $F_x(x)$ a probabilidade de ocorrência de um evento menor ou igual a x e é dada pela função de distribuição acumulada de probabilidade é representada pela função Gama incompleta, segundo THOM (1958):

$$F(x_0) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{x_0} u^{\alpha-1} e^{-\frac{u}{\beta}} du \quad (3.12)$$

em que,

$F(x_0)$ = probabilidade de ocorrer um valor menor ou igual a x;

x = variável aleatória contínua, positiva e independente;

$\Gamma(\alpha)$ = função matemática 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. O principal método para estimar seus parâmetros é o método de máxima verossimilhança; que para satisfazer a condição $\alpha > 0$ (por definição) e após algumas considerações matemáticas na equação original (Equação 3.12) a solução que interessa é (HAAN, 1977):

$$\hat{\alpha} = \frac{1 + \sqrt{1 + 4(\ln(\bar{x}) - x_g) / 3}}{4(\ln(\bar{x}) - x_g)} \quad (3.13)$$

O estimador do parâmetro β , pode ser obtido por:

$$\beta = \frac{\bar{x}}{\hat{\alpha}} \quad (3.14)$$

sendo \bar{x} a média aritmética e x_g a média geométrica da amostra. A função de distribuição acumulada da distribuição de probabilidade Gama 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 $t = \frac{u}{\beta} \rightarrow du = \beta dt$, não mudando o limite inferior zero e

substituindo x por t , e após as simplificações adequadas, tem-se, da Equação 3.15:

$$F_x(t) = \frac{1}{\Gamma(\alpha)} \int_0^t t^{\alpha-1} e^{-t} dt \quad (3.15)$$

Considerando o desenvolvimento em série de e^{-t} , multiplicando-se e dividindo-se a equação anterior por e^{-t} , reunindo, do produto resultante, os termos em t, t^2, t^3, \dots e os demais termos correspondentes aos diversos expoentes, tem-se, segundo ASSIS et al. (1996):

$$F_x(t) = \frac{t^\alpha}{\alpha \Gamma(\alpha) e^t} \left[1 + \frac{t}{\alpha + 1} + \frac{t^2}{(\alpha + 1)(\alpha + 2)} + \frac{t^3}{(\alpha + 1)(\alpha + 2)(\alpha + 3)} + \dots \right] \quad (3.16)$$

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

Distribuição de probabilidade GEV

A distribuição GEV (do inglês “Generalized Extreme Value”) foi introduzida por JENKINSON (1955), incorporando as 3 formas assintóticas: Gumbel (Tipo I), Fréchet (Tipo II) e Weibull (Tipo III). Sua FDP é dada por:

$$f(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\left(\frac{1+\xi}{\xi} \right)} \cdot \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \quad (3.17)$$

em que ζ , σ e μ são, respectivamente, os parâmetros de forma, escala e locação. Se ζ for negativo, a GEV representa a forma assintótica de valores mínimos (Tipo III) e existe apenas para $x < \frac{(\mu - \sigma)}{\xi}$. Se ζ for positivo, a GEV representa uma distribuição Tipo II (Fréchet), definida para $x > \frac{(\mu - \sigma)}{\xi}$. Se $\xi = 0$, tem-se a Distribuição Gumbel. Sua FCP é dada por:

$$F_x(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \quad (3.18)$$

A distribuição GEV apresenta 3 momentos estatísticos:

$$E[x] = \bar{x} = \mu + \frac{\sigma}{\xi} [\Gamma(1 - \xi) - 1] \quad (3.19)$$

$$Var(x) = \left(\frac{\sigma}{\xi} \right)^2 \cdot [\Gamma(1 - 2 \cdot \xi) - \Gamma^2(1 - \xi)] \quad (3.20)$$

$$\gamma = (\text{ sinal de } \xi) \cdot \left[\frac{\Gamma(1 - 3 \cdot \xi) + 3 \cdot \Gamma(1 - \xi) \cdot \Gamma(1 - 2 \cdot \xi) - 2 \cdot \Gamma^3(1 - \xi)}{[\Gamma(1 - 2 \cdot \xi) - \Gamma^2(1 - \xi)]^2} \right] \quad (3.21)$$

Para estimar os parâmetros da distribuição GEV deve-se inicialmente calcular o parâmetro de forma pela Equação (3.21). Para isto, é importante observar qual sinal ζ tem para a situação em estudo, o que é obtido mediante análise do coeficiente de assimetria

(distorção). Para $\zeta = 0$, $\gamma = 1,1396$; valores de γ superiores a este valor, $\zeta < 0$; valores de γ menores que 1,1396 implicamos em $\zeta > 0$ até o valor de $-\frac{1}{3}$. A estimativa de uma variável hidrológica associada a um tempo de retorno é dada por, considerando-se frequência de excedência:

$$x_{TR} = \mu + \frac{\sigma}{\zeta} \left\{ 1 - \left[-LN \left(1 - \frac{1}{TR} \right) \right]^{-\zeta} \right\} \quad (3.22)$$

Distribuição de probabilidade Gumbel

A função densidade de probabilidade (fdp) Gumbel é dada por:

$$f_x(x) = \alpha \times e^{\{-\alpha(x-\mu) - e^{-\alpha(x-\mu)}\}} \quad (3.23)$$

A integração da $f_x(x)$ fornece a função de distribuição acumulada:

$$P(X \leq x_i) = e^{-e^{-\alpha(x-\mu)}} \quad (3.24)$$

Esta distribuição apresenta os dois primeiros parâmetros de uma distribuição de probabilidades, ou seja, a média (μ) e a variância (σ), que são calculados pelas expressões abaixo, considerando se o método dos momentos:

$$\hat{\alpha} = \frac{1,2826}{s} \quad (3.25)$$

$$\hat{\mu} = \bar{x} - 0,45 \times s \quad (3.26)$$

em que \bar{x} e s correspondem, respectivamente, à média e o desvio padrão da série histórica.

Para estimativa de uma variável hidrológica x em função do TR, aplica-se a equação abaixo, fruto da manipulação da Equação (3.24) e da consideração de TR como função da probabilidade de excedência:

$$x_{TR} = \frac{-LN \left(-LN \left(1 - \frac{1}{TR} \right) \right)}{\alpha} + \mu \quad (3.27)$$

Distribuição de probabilidade Log-normal

Uma outra distribuição, testada por Huf e Neil (1959), num trabalho de comparação entre vários métodos para analisar frequência de precipitação, é a distribuição log-normal, a qual assume que os logaritmos das variáveis aleatórias são normalmente distribuídos. Conforme Mirshawka (1971), a função de densidade acumulada Log-normal com dois parâmetros e a três parâmetros, são representadas pela seguinte equação:

$$f_x(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x-a)-\mu]^2}{2\sigma^2}\right) \quad (3.28)$$

em que,

$f(x)$ = probabilidade da variável ser menor ou igual a x ;

e = base do logaritmo neperiano;

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:

$$F(a \leq x \leq b) = \int_a^b \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x-a)-\mu]^2}{2\sigma^2}\right) dx \quad (3.29)$$

O valor de “ a ” pode ser zero, quando se considera a distribuição log-normal com dois parâmetros, ou um valor mínimo da série, quando se considera log-normal com 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).

Distribuição de probabilidade logística

A distribuição Logística apresenta normalmente duas expressões. Uma denominada de fórmula geral e outra de forma padrão. A distribuição Logística é derivada do trabalho de Verhulst, Professor de Análise na Faculdade Militar Belga. Ele a utilizou para modelar o crescimento da população na Bélgica no início de 1800.

A expressão geral da função de densidade de probabilidade (fdp) Logística é dada por:

$$f_x(x) = \frac{\beta^{-1} e^{(x-\alpha)/\beta}}{\left[1 + e^{(x-\alpha)/\beta}\right]^2} \quad \text{para } x \in R, \quad \beta > 0 \quad (3.30)$$

ou

$$f_x(y) = \frac{(1/\beta)e^y}{\left[1 + e^y\right]^2}, \quad \text{para } x \in R, \quad y = \frac{x-\alpha}{\beta} \quad (3.31)$$

onde α é o parâmetros de locação e β é o parâmetro de escala.

A função densidade de probabilidade Logística padrão é dada por:

$$f(x) = \frac{e^x}{\left[1 + e^x\right]^2}, \quad \text{para } x \in R \quad (3.32)$$

ou

$$f(x) = \frac{e^x}{\left[1 + e^{-x}\right]^2}, \quad \text{para } x \in R \quad (3.33)$$

Distribuição de probabilidade 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:

$$f_x(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]; & \text{para } x \geq 0 \\ 0; & \text{para } x < 0 \end{cases} \quad (3.34)$$

em que μ é a média e σ o desvio-padrão populacional.

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

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \quad (3.35)$$

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 $N(0,1)$, isto é, distribuição normal com média zero e variância um. A variável Z é chamada variável reduzida, e a Equação (3.36) toma a seguinte forma:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{z^2}{2}\right) dz \quad (3.36)$$

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:

$$F(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5) \quad (3.37)$$

para $z \leq 0$, já que para $z < 0$ tem-se $F(z) = 1 - F(-z)$, em que:

$$t = \frac{1}{1 + kz}$$

$$k = +0,2316419;$$

$$a_1 = +0,31938153;$$

$$a_2 = -0,356563782;$$

$$a_3 = +1,781477937;$$

$$a_4 = -1,821255978;$$

$$a_5 = +1,330274429; \text{ e}$$

$z =$ variável reduzida ou transformada de x .

Distribuição de probabilidade Weibull

A distribuição de probabilidade Weibull foi proposta primeiramente por Fischer e Tippett, em 1928, tendo sido desenvolvida independentemente por Walodi Weibull, físico sueco em 1939. 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:

$$f_x(x) = \frac{\gamma}{\beta} \left(\frac{x - \alpha}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x - \alpha}{\beta}\right)^\gamma\right], \quad \text{para } x \geq \alpha \quad (3.38)$$

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 com três parâmetros. A sua função de distribuição acumulada é:

$$F(x) = \int_0^{\infty} f(x)dx = 1 - \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^{\gamma} \right] \quad (3.39)$$

O parâmetro α é chamado de parâmetro de posição, pois controla a posição da curva sobre o eixo das abscissas. O parâmetro de escala β controla as dimensões que a curva assume, dada uma forma constante. O parâmetro γ , é chamado parâmetro de forma. 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. O principal método de ajuste da distribuição Weibull (COUTO, 1980) e o da máxima verossimilhança, que consiste em determinar os valores de γ e β pela suas equações fundamentais. Nota-se porém que β é função de γ , enquanto este é função de β e de si próprio.

Normalmente, a solução desse tipo de sistema (Equação 3.40) é obtido resolvendo a equação por um processo iterativo, por exemplo, o método da bissecção, sendo β encontrado, a partir de γ aproximado, e introduzido na Equação (3.41).

$$\frac{\sum_{i=1}^n x_i^{\hat{\gamma}} \ln(x_i)}{\sum_{i=1}^n x_i} - \frac{\sum_{i=1}^n \ln(x_i)}{n} - \frac{1}{\hat{\gamma}} = 0 \quad (3.40)$$

$$\hat{\beta} = \left(\frac{\sum_{i=1}^n \ln(x_i^{\hat{\gamma}})}{n} \right)^{\frac{1}{\hat{\gamma}}} \quad (3.41)$$

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.

Adequação estatística de uma distribuição de probabilidades

Para a aplicação de uma distribuição de probabilidades é indispensável analisar se a mesma representa adequadamente bem a relação funcional entre os valores do evento e as respectivas frequências amostrais dos mesmos. Para isto, há necessidade de se comprovar

previamente se a distribuição é adequada para a série histórica a ser trabalhada. A comprovação é feita com base em testes estatísticos não paramétricos, os quais, são conhecidos como Testes de Aderência Estatística.

Teste de Kolmogorov-Smirnov

Neste teste, promove-se o cálculo da diferença entre as frequências observadas (amostrais) e as frequências esperadas com base na distribuição de probabilidades, comparando-se a maior diferença obtida a um valor que correspondente à estatística do teste. Esta estatística é obtida em função do tamanho da amostra (n) e nível de significância (α) a ser adotado (5% na maioria das vezes). A hipótese de nulidade a ser testada é a hipótese H_0 de que a frequência observada poderá ser estimada pela distribuição de probabilidades, ou seja, como o valor tabelado é estatisticamente nulo, pode se concluir que valores menores ou iguais a este serão também estatisticamente nulos. Desta forma, tem-se:

$$|\Delta F|_{\text{calculado máx}} \leq |\Delta F|_{\text{tabelado}(n,\alpha)} \quad (3.42)$$

Nesta situação, a distribuição de probabilidades será adequada, pois $|\Delta F|_{\text{calculado máx}}$ será nulo estatisticamente e, portanto, a frequência observada não difere da esperada. Observa-se que apenas a máxima diferença entre as frequências é considerada neste teste. Desta forma, o Teste de Kolmogorov-Smirnov é inteiramente qualitativo, significando que o mesmo permite apenas a conclusão de que a distribuição de probabilidades é adequada ou não, não havendo embasamento suficiente para se concluir a respeito da precisão e comparação entre distribuições distintas. Se na Equação (3.41) ocorrer o contrário, a distribuição não será adequada, devendo-se buscar o ajuste de outra.

Teste de Qui-quadrado

Este teste é mais rigoroso que o anterior por agrupar os dados da série histórica em classes de frequência e acumular os erros entre as frequências observada e teórica, com participação de todas as classes e não apenas a máxima diferença. A soma destes erros (obtida pela soma dos erros de todas as classes) gera o valor de χ^2 calculado. A estatística do teste é obtida por meio da tabela de χ^2 adotando-se o valor tabelado com base em graus de liberdade da distribuição e nível de significância. Para que a distribuição de probabilidades tenha

aderência aos dados, o valor de χ^2 calculado deve ser menor que o valor de χ^2 tabelado. Assim, tem-se:

$$\chi^2_{\text{calculado}} = \sum_{i=1}^n \frac{(f_{\text{obsi}} - f_{\text{teóricoi}})^2}{f_{\text{teórico}}} \quad (3.43)$$

em que n é o número de classes, f_{obsi} e $f_{\text{teóricoi}}$ são, respectivamente, as frequências observada e teórica na classe i . As classes com menos de três valores devem ser agrupadas com as classes vizinhas, seguindo os critérios de aplicação do teste. Os Graus de Liberdade a serem adotados neste teste podem ser obtidos considerando-se uma situação intermediária entre o número de classes menos 1 e o número de classes menos número de parâmetros da distribuição menos 1. Um detalhe adicional é de que a Equação (3.43) representa uma forma de cálculo do quadrado médio do erro e todas as frequências participam do mesmo. Desta forma, Walpole e Myers (1978) consideram o teste de χ^2 um cálculo de precisão do ajuste da distribuição de probabilidades.

Estudo complementar

A distribuição normal bivariada

Quanto à distribuição normal bivariada, trata-se de uma distribuição normal com duas dimensões das variáveis contínuas X_1 (no caso, a duração da seca) e X_2 (a severidade da seca).

Portanto, é uma generalização da distribuição normal para uma variável X com função densidade de probabilidade denotada por:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{t}{2}\right) \quad (3.44)$$

$$t = \left\{ \frac{1}{(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \right\} \quad (3.45)$$

em que tem-se $(-\infty < x_1, x_2 < \infty)$.

Com parâmetros média μ , variância σ^2 e a correlação ρ entre as variáveis:

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad (3.46)$$

em que μ é o vetor de médias, sendo μ_1 a média da duração de seca e μ_2 é a média da severidade da seca.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & Cov(\sigma_1, \sigma_2) \\ Cov(\sigma_1, \sigma_2) & \sigma_2^2 \end{bmatrix} \quad (3.47)$$

em que Σ é a matriz de covariâncias: σ_1^2 é a variância da duração da seca; σ_2^2 é a variância da severidade da seca e $Cov(\sigma_1, \sigma_2)$ é a covariância entre as duas variáveis (MORRISON, 1990).

As propriedades da função distribuição de probabilidade normal bivariada são as seguintes:

- $f(x_1, x_2)$ é a função densidade de probabilidade;
- a variável x_1 possui distribuição normal (μ_1, σ_1^2) e a variável x_2 possui distribuição normal (μ_2, σ_2^2) ;
- ρ é o coeficiente de correlação entre x_1 e x_2 .

Portanto, as variáveis reduzidas z_1 e z_2 bem como a função densidade de probabilidade da distribuição normal bivariada $f(z_1, z_2)$ são assim representadas:

$$z_1 = \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \quad (3.48)$$

$$z_2 = \left(\frac{x_2 - \mu_2}{\sigma_2} \right) \quad (3.49)$$

$$f(z_1, z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2} \frac{1}{(1-\rho^2)} (z_1^2 - z_1\rho z_2 + z_2^2) \right] \quad (3.50)$$

A Figura 3.4, ilustra a distribuição normal bivariada para duas variáveis:

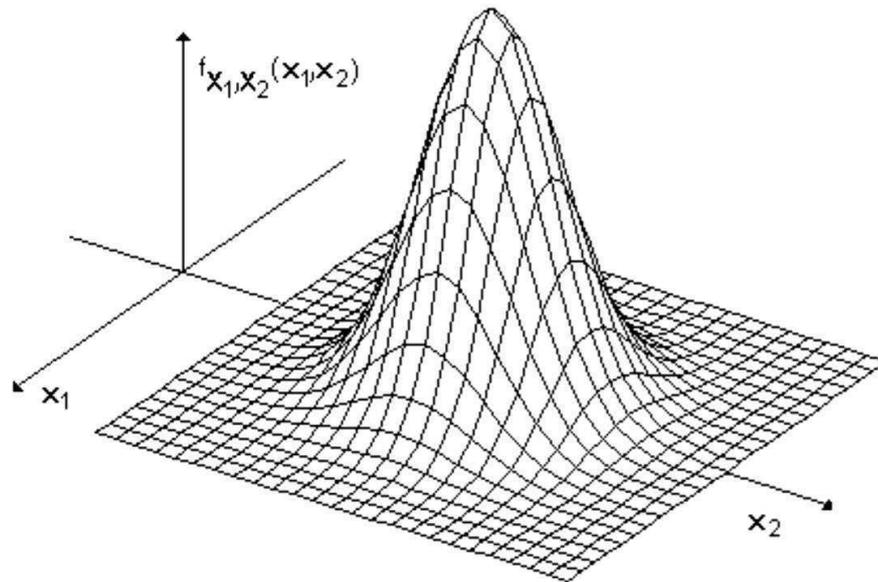


Figura 3.1: Representação da normal bivariada

Verificação da normalidade bivariada

Inicialmente supõe-se que as variáveis em estudo seguem uma distribuição de probabilidades normal bivariada, mas é necessário verificar essa normalidade. Na literatura existem muitos métodos analíticos propostos para realizar essa verificação no entanto, decidir qual método usar é um processo desafiador, uma vez que cada método pode dar resultados diferentes em dadas condições. Assim, pode-se dizer que o melhor método é aquele que é válido em qualquer condição. Além de resultados numéricos, é muito útil usar métodos gráficos para decidir sobre normalidade bivariada. Combinar os resultados numéricos a partir de vários métodos, com abordagens gráficas pode ser útil e fornecer decisões mais confiáveis.

Para a verificação da suposição de normalidade bivariada foi utilizado o pacote MVN versão 3.9 (KORKMAZ et al., 2015), no software gratuito R. Para este estudo foi adotado um nível de significância de 0,05, neste utilizou-se do teste de Henze-Zirkler e incluiu-se gráficos qui-quadrado Q-Q.

Teste MVN de Henze-Zirkler

O teste de Henze-Zirkler baseia-se numa distância funcional não negativa, que mede a distância entre duas funções de distribuição. Se os dados são distribuídos como bivariado, e a estatística normal de teste é aproximadamente log-normalmente distribuídos. Primeiro, a média, variância e parâmetro de suavidade são calculado. Em seguida, a média e a variância

são log-normalizado e o valor p é estimada. A estatística de teste de teste de normalidade bivariada de Henze-Zirkler é dado na equação:

$$HZ = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n e^{-\frac{\beta^2}{2} D_{ij}} - 2(1 + \beta^2)^{-\frac{p}{2}} \sum_{i=1}^n e^{-\frac{\beta^2}{2(1+\beta^2)} D_i} + n(1 + 2\beta^2)^{-\frac{p}{2}} \quad (3.51)$$

onde:

p é o número de variáveis

$$\beta = \frac{1}{\sqrt{2}} \left(\frac{n(2p+1)}{4} \right)^{\frac{1}{p+4}}$$

$$D_{ij} = (x_i - x_j) S^{-1} (x_i - x_j)$$

$$D_i = (x_i - \bar{x}) S^{-1} (x_i - \bar{x}) = m_{ii}$$

A partir da Equação (3.51), D_i é a distância de Mahalanobis ao quadrado i^{th} observa-se o centroide D_{ij} com a distância de Mahalanobis entre i^{th} e j^{th} . Se os dados seguirem uma distribuição normal bivariada, a estatística teste (HZ) é aproximadamente uma distribuição log-normal com média μ e variância σ^2 , visto abaixo:

$$\mu = 1 - \frac{a^{-\frac{p}{2}} \left(1 + p\beta^{\frac{2}{a}} + (p(p+2)\beta^4) \right)}{2a^2} \quad (3.53)$$

$$\sigma^2 = 2(1+4\beta^2)^{-\frac{p}{2}} + \frac{2a^{-p}(1+2p\beta^4)}{a^2} + \frac{3p(p+2)\beta^8}{4a^4} - 4w_\beta^{-\frac{p}{2}} \left(1 + \frac{3p\beta^4}{2w_\beta} + \frac{p(p+2)\beta^8}{2w_\beta^2} \right) \quad (3.54)$$

em que $a = 1 + 2\beta^2$ e $w_\beta = (1 + \beta^2)(1 + 3\beta^2)$. Assim, a média e a variância do log-normal HZ e a estatística pode ser definida como:

$$\log(\mu) = \log \left(\sqrt{\frac{\mu^4}{\sigma^2 + \mu^2}} \right) \quad (3.56)$$

e

$$\log(\sigma^2) = \log \left(\frac{\sigma^2 + \mu^2}{\sigma^2} \right) \quad (3.57)$$

Usando parâmetros da distribuição log-normal μ e σ , pode-se testar a significância na normalidade bivariada. O teste estatístico **Wald** para a normalidade é dada pela Equação (3.58):

$$z = \frac{\log(HZ) - \log(\mu)}{\log(\sigma)} \quad (3.58)$$

4 RESULTADOS E DISCUSSÃO

Análise climática

Na análise do comportamento climático do Semiárido brasileiro (SAB) utilizaram-se dados de totais mensais precipitados coletados em 320 postos pluviométricos, especificando uma série temporal de 1984 a 2014, gentilmente cedidos pela Agência Nacional de Águas (ANA). Utilizou-se também do software Surfer 8.0 para efetuar o mapa de espacialização dos postos. Na Figura 4.1 pode ser visto a distribuição espacial dos postos pluviométricos na área de estudo. Posteriormente, esses dados foram tratados e organizados para se obter as estatísticas básicas para cada um desses postos.

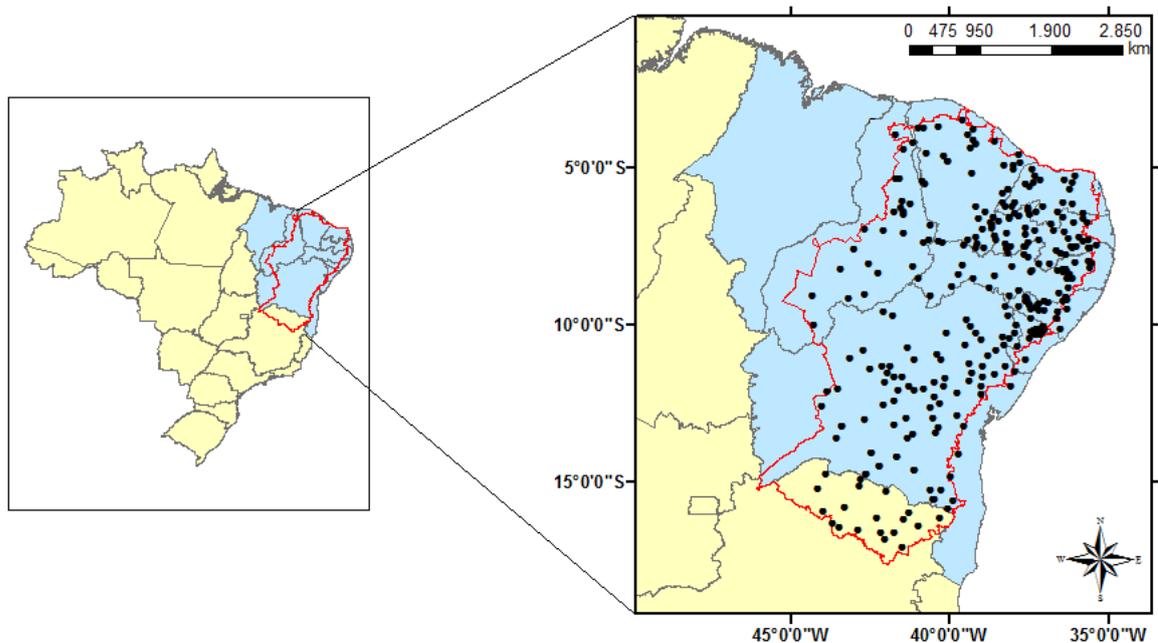


Figura 4.1: Distribuição espacial dos postos pluviométricos na área de estudo

Com base nos totais mensais precipitados dos postos pluviométricos foram obtidas as estatísticas descritivas da precipitação pluvial do semiárido brasileiro, no período estudado. Observa-se na Tabela 4.1 que há uma predominância do período chuvoso no verão, entre os meses de dezembro e maio, destacando-se os meses de fevereiro, março e abril (FMA) que representam o trimestre mais chuvoso do SAB, essa variabilidade interanual da pluviometria na SAB está associada às variações de padrões de Temperatura da Superfície do Mar (TSM)

sobre os oceanos tropicais, os quais afetam a posição e a intensidade da Zona de Convergência Intertropical (ZCIT) sobre o Oceano Atlântico (HASTENRATH, 1984; citados por NOBRE E MELO, 2001).

Os resultados da análise descritiva das médias mensais da precipitação visto na Tabela 4.1 tem-se que todos os meses apresentaram coeficiente de assimetria positivo, com assimetria à direita ($CA > 0$), outra observação importante é quanto aos coeficientes de assimetria é que esse fato sugere que os dados de precipitação podem se ajustar à uma função distribuição de probabilidades de Gumbel, Log-normal, Gama, dentre outras com assimetria positiva. Os coeficientes de variação ($CV > 20\%$) indicam que há alta dispersão entre os valores médios de precipitação, ou seja, há grande variabilidade da chuva na região do Semiárido brasileiro. Todos os valores médios são maiores do que as medianas, isso confirma a assimetria positiva dos dados. Com relação à curtose, todos os valores médios mensais apresentaram coeficientes de curtose ($CC > 0$), indicando que a distribuição dos dados é do tipo leptocúrtica, ou seja, significa que é mais fácil obter valores que se afastam da média, isso acontece devido nesses meses dar-se início a estação chuvosa da maior parte da região de estudo.

Tabela 4.1 - Estatística descritiva dos dados de precipitação do Semiárido brasileiro no período de 1984 à 2014.

Meses	Mínimo	Máximo	1ºQuartil	Mediana	3ºQuartil	Média	DP	CV (%)	CA	CC
Janeiro	17,1	408,2	55,6	86,3	115,7	90,2	46,6	51,7	1,4	5,8
Fevereiro	20	428,2	58,7	84,2	113,9	92,1	48,5	52,6	2,1	8,6
Março	25,1	456,6	74,1	113,2	167	127,9	68	53,2	1,3	2,8
Abril	17,9	425,3	55,6	90,4	147,7	107,2	66,4	62	1,3	2
Maio	1,9	257,3	29,3	63,6	97,9	67,9	45,3	66,7	0,6	0,1
Junho	0	211,5	11,6	38,1	59,3	47,3	45,7	96,7	1,3	1,3
Julho	0	210,5	7,7	21,7	53,8	39,2	44,5	113,7	1,6	1,9
Agosto	0	125,5	3,3	10,5	35	23,9	28,4	119	1,5	1,6
Setembro	0	86,9	3,4	9,8	22,4	15,9	17,2	108,1	1,6	2,5
Outubro	0	198,3	9,4	19,5	38,4	26	23,2	89,2	2	8,7
Novembro	0,4	244,7	10,5	25	79,2	49,6	52,9	106,7	1,3	0,7
Dezembro	1,6	291,3	26,1	44,2	89,2	63,5	53,4	0,8	1,6	2,3

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

Delimitação de sub-regiões pluviometricamente homogêneas

Depois de observar o comportamento da precipitação no SAB, utilizou-se o método de Análise de Agrupamento para delimitar a região de estudo em três sub-regiões pluviometricamente homogêneas e distintas entre si. A Figura 4.2 exibe a delimitação das três sub-regiões, denominadas de sub-região 1, sub-região 2 e sub-região 3.

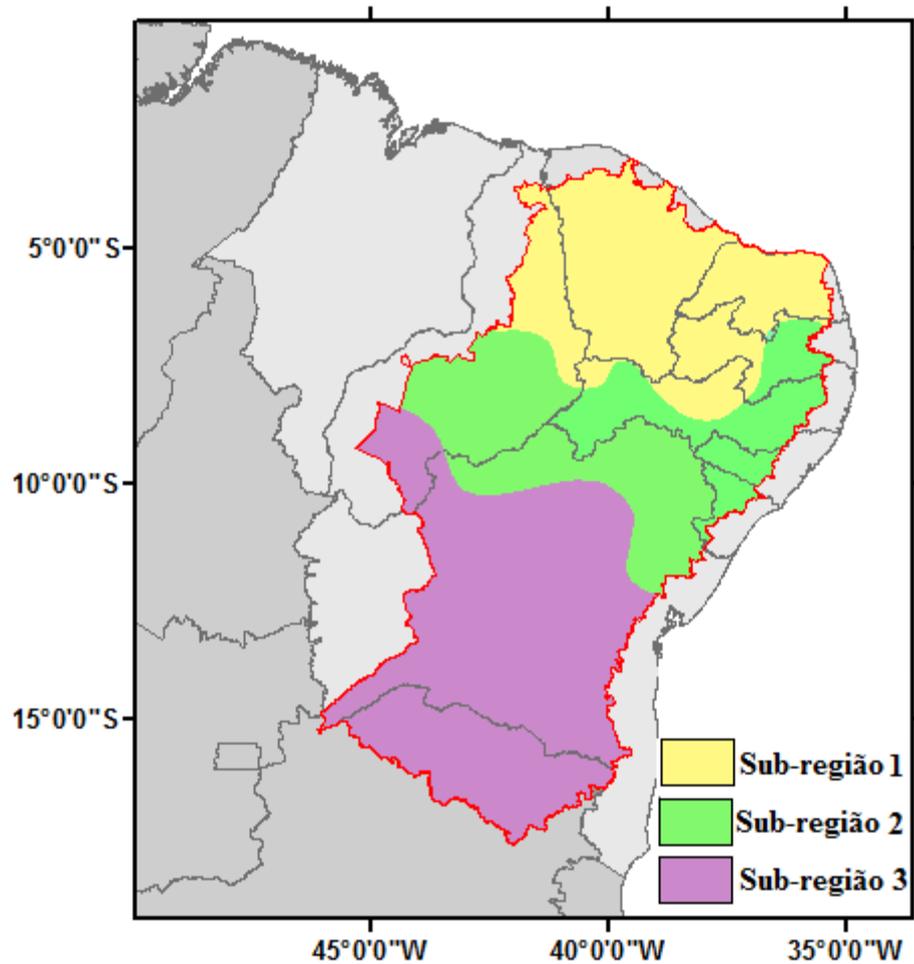


Figura 4.2: Delimitação do Semiárido brasileiro (SAB) em três sub-regiões

Para cada sub-região foi realizado um estudo estatístico através das medidas descritivas (total máximo, total mínimo e média) da precipitação pluviométrica. A sub-região 1 contém 82 postos pluviométricos, ou seja, quase 26% dos postos da área de estudo. A precipitação média mensal é igual a 76,2 mm, a máxima é igual a 230,0 mm e a mínima de 2,9 mm. Na Figura 4.3 pode ser visto que o trimestre mais chuvoso é composto pelos meses

de fevereiro, março e abril. Nessa sub-região 1 as chuvas, nos meses mais chuvosos, variam de 150,0 mm a 230,0 mm. Os meses de agosto a setembro são os mais secos dessa sub-região.

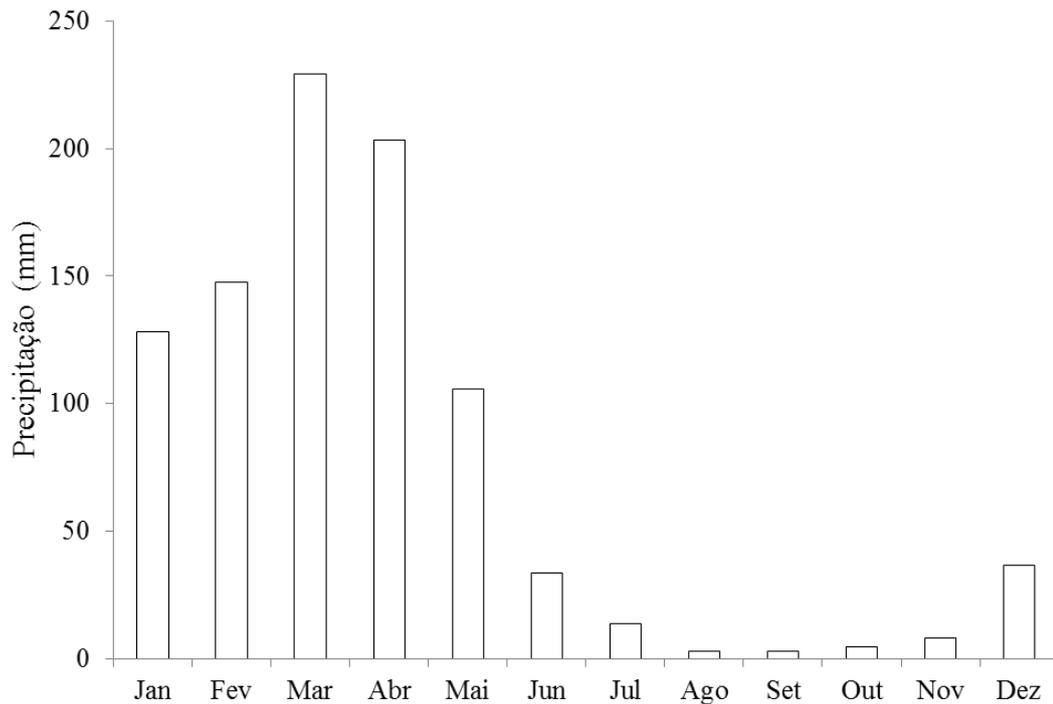


Figura 4.3: Precipitação média mensal da sub-região 1

A sub-região 2 contém 154 postos pluviométricos, sendo a com maior número de postos pluviométricos da área de estudo, com cerca de 48% dos postos, apresenta precipitação média mensal igual a 54,9 mm, máxima de 95,3 mm e mínima de 20,3 mm. Na Figura 4.4 pode ser visto que o trimestre mais chuvoso está entre os meses de março e maio, com chuvas variando entre 90,0 mm e 95,0 mm. Destaque-se nesta sub-região 2 uma melhor distribuição temporal da precipitação pluvial.

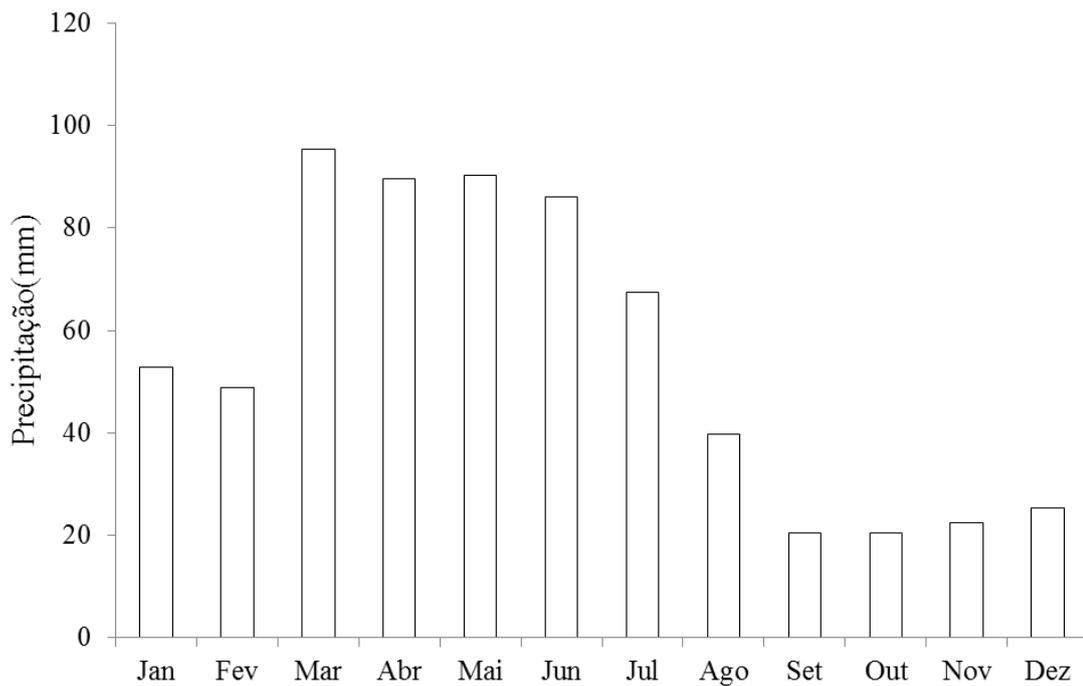


Figura 4.4: Precipitação média mensal da sub-região 2

A sub-região 3 contém 84 postos pluviométricos, cerca de 26% dos postos da região em estudo. A Figura 4.5 exibe os meses de novembro a janeiro como o trimestre mais chuvoso. As chuvas dessa sub-região 3 variam entre 100,0 mm e 137,0 mm, com média de 56,6 mm. A precipitação máxima é igual a 137,4 mm e a mínima é igual a 3,0 mm. Nessa sub-região os meses mais secos estão entre maio e setembro.

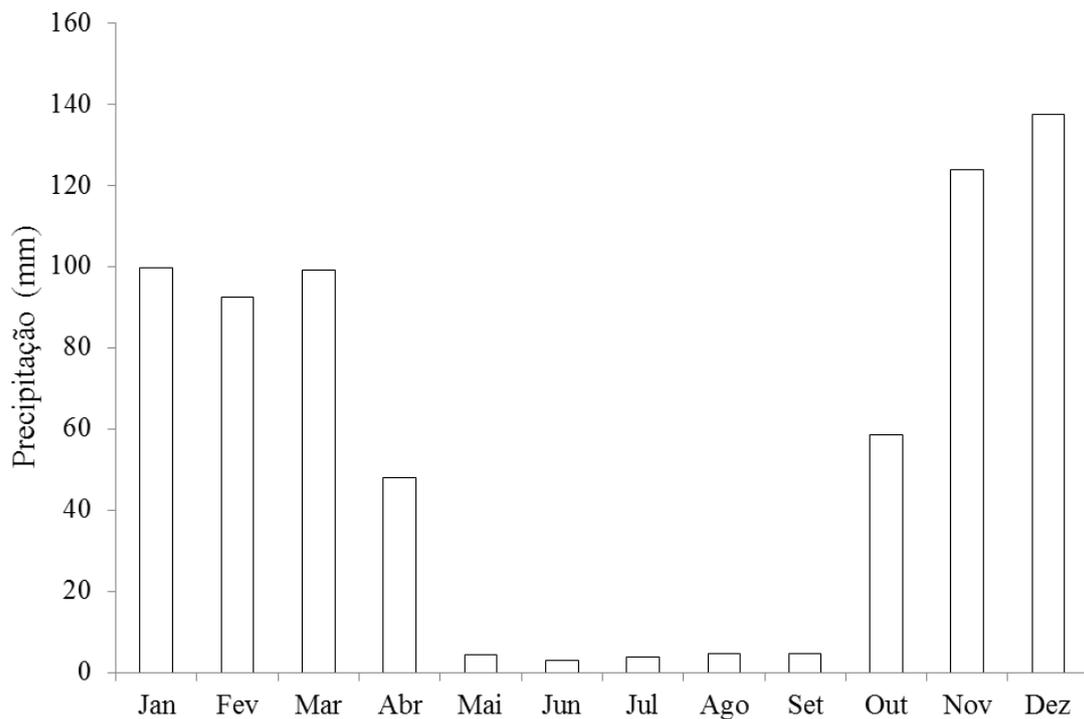


Figura 4.5: Precipitação média mensal da sub-região 3

Delimitação de sub-regiões homogêneas com base na variável Severidade de secas

Após caracterizar o comportamento médio da precipitação pluvial da área de estudo, o SAB, o foco principal deste trabalho foi o de delimitar sub-regiões homogêneas com base nas variáveis Severidade e Duração de secas. Em seguida, foram ajustadas funções distribuição de probabilidades marginal a essas variáveis, no âmbito do SAB, nos últimos 30 anos. Com base nas anomalias das chuvas e método (“run”) foram estimados os valores das variáveis Severidade e Duração de secas em todo o SAB no período de estudo. Como já efetuado na primeira parte deste estudo utilizou-se a Análise de Agrupamento para delimitar três sub-regiões homogêneas com base nas variáveis Severidade e Duração de secas. A Figura 4.6 exibe essas sub-regiões, distintas entre si, com base na variável Severidade de seca.

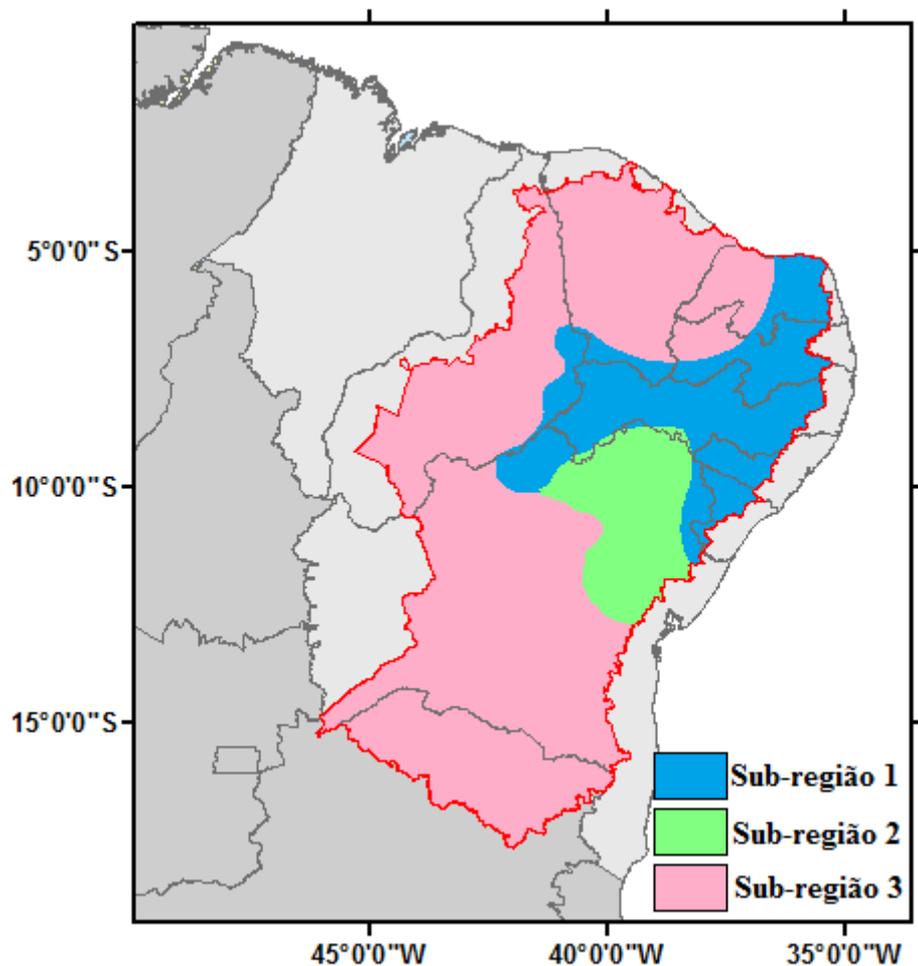


Figura 4.6: Sub-regiões homogêneas com base nos índices de Severidade de Seca (mm)

A sub-região 1 ocupa o leste do Estado do Rio Grande do Norte até o leste do Estado de Sergipe, estendendo-se para o extremo norte do Estado da Bahia, sul do Estado do Ceará e leste do Estado do Piauí. As regiões semiáridas dos Estados de Sergipe e Alagoas estão 100% contidas nessa sub-região. Já o semiárido do Estado de Pernambuco está quase que totalmente contido nessa sub-região. A sub-região 2 é bem menor que as demais e localiza-se no leste e norte do Estado da Bahia e sul do Estado de Pernambuco. Por fim, a sub-região 3 se destaca por ser a maior delas. Abrange o oeste da Paraíba e do Rio Grande do Norte, quase todo o Estado do Ceará, do Piauí e da Bahia, além da parte semiárida do Estado de Minas Gerais.

A Tabela 4.2 exhibe a estatística descritiva das sub-regiões, onde se pode ver que a sub-região 3 apresenta as maiores variações, com média de 260,2mm e um máximo de Severidade de Seca de 821,8mm. Essa sub-região é também a maior em número de postos, apresenta-se

mais seca em relação às outras. Já na sub-região 2, a Severidade da seca se apresenta menos intensa. Observa-se que o coeficiente assimetria se apresenta sempre positivo nas sub-regiões 1 e 3, indicando que as duas sub-regiões têm distribuição assimetricamente positiva; já a sub-região 2 apresenta assimetria negativa. Além desse fato também é importante observar que os valores dos coeficientes de curtose são todos maiores que zero nas três sub-regiões. Outro índice importante na Tabela 4.2 é a Intensidade da seca, dado pela razão entre Severidade e Duração de seca. Nota-se que a área que apresentou maior Intensidade de seca foi a sub-região 3, a maior das sub-regiões. Esse resultado corrobora com a continentalidade do fenômeno das secas, agrava-se à medida que se adentra no continente.

Tabela 4.2 - Estatísticas descritivas das sub-regiões para a Severidade de seca (mm).

Estatística	Sub-região 1	Sub-região 2	Sub-região 3
Severidade média	179,54	78,46	260,25
Desvio padrão	53,55	15,3	117,57
Severidade mínima	30,04	29,65	73,71
Severidade máxima	380,86	101,96	821,75
Coeficiente de assimetria	0,58	-0,87	1,56
Coeficiente de curtose	1,44	1,27	4,56
Número de observações	83	32	204
Intensidade de seca (S/D)	39,0	24,7	48,5

Fonte: Elaborada pelo autor

A Figura 4.7 exibe o mapa da variabilidade espacial da Severidade de seca (mm) no SAB. Pode-se notar que os Estados do Piauí, Ceará, oeste da Paraíba, oeste do Rio Grande do Norte, oeste e sul da Bahia e norte de Minas Gerais foram as áreas mais afetadas pelas secas no período estudado. Esse fato foi evidenciado na análise descritiva dos dados, Tabela 4.1. Nota-se, também, na Figura 4.7 o efeito da continentalidade da Severidade de seca, com exceção da parte setentrional da região de estudo.

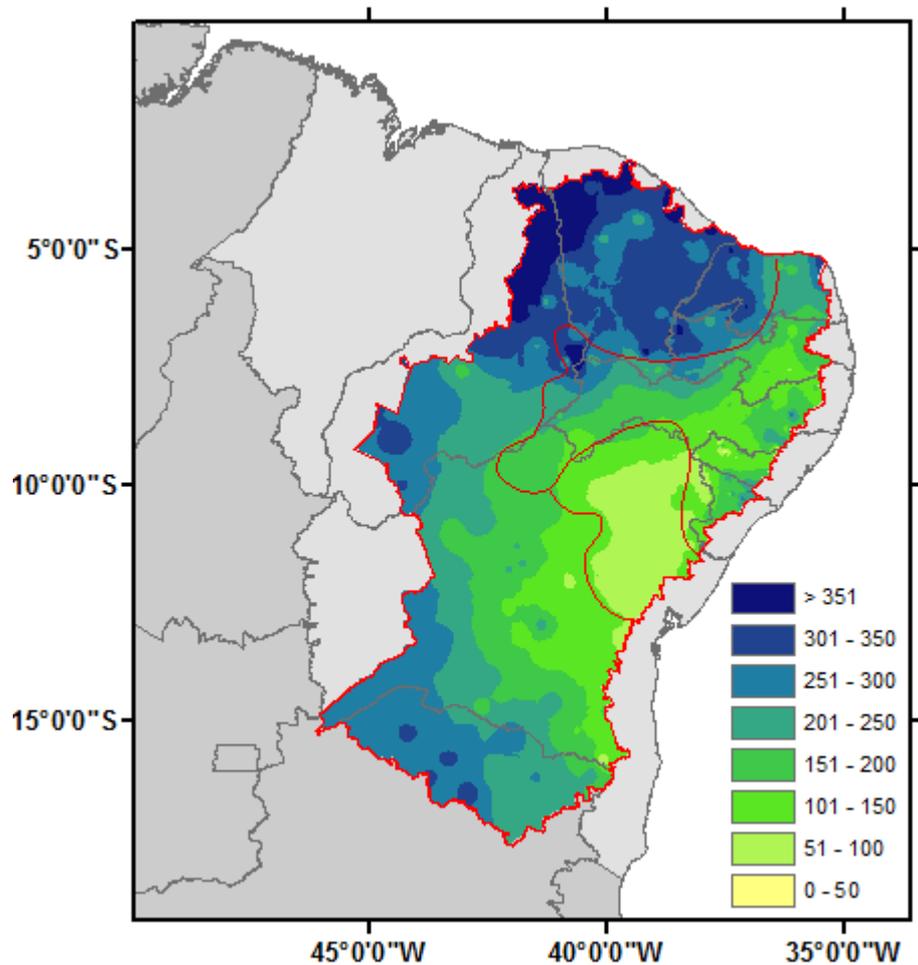


Figura 4.7: Variabilidade espacial da Severidade de seca no SAB

Delimitação de sub-regiões homogêneas com base na variável Duração de Secas

A Figura 4.8 exibe a variabilidade da Duração de secas (meses) no SAB. Pode-se notar que grande parte do Estado do Ceará, norte da Paraíba, centro-oeste do Rio Grande do Norte e oeste de Pernambuco foram as áreas mais afetadas quanto à Duração de secas. Tal como já detectado pela Severidade de secas, a Duração também registra o efeito da continentalidade, ou seja, aumenta à medida que se adentra no continente, com exceção do leste do Rio Grande do Norte e alguns áreas isoladas.

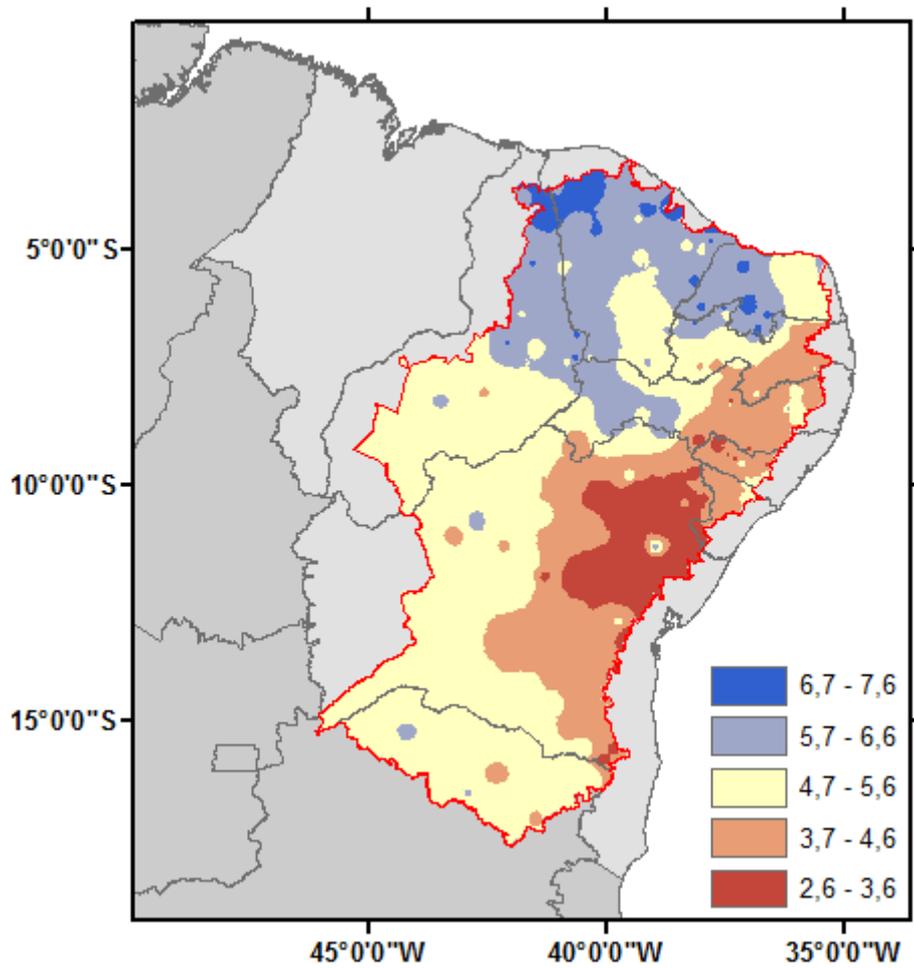


Figura 4.8: Variabilidade espacial da Duração de secas em meses no SAB

A Tabela 4.3 exibe a estatística descritiva das sub-regiões, com relação à Duração de secas (meses). Pode ser visto que a sub-região 2 apresenta as maiores durações de secas, com média de 6,63 meses, máxima Duração de secas de 7,64 meses e mínima Duração de 5,45 meses. Os coeficientes de assimetria (C.A) e curtose (C.C) são moderadamente assimétricos e quase mesocúrticos.

Tabela 4.3 - Estatísticas descritivas das sub-regiões para a Duração de secas (meses)

Estatística	Sub-região 1	Sub-região 2	Sub-região 3
Duração média	4,67	3,22	5,17
Desvio padrão	0,86	0,79	1,01
Duração mínima	2,9	2,37	2,79
Duração máxima	7,2	6,25	7,47
CA	0,7	1,97	0,16
CC	0,14	4,61	0,32
CV (%)	0,18	0,25	0,20

Fonte: Elaborada pelo autor

Ajuste de funções distribuição de probabilidades marginal à variável Severidade de secas

Após analisar o comportamento da Severidade de secas nas três sub-regiões, utilizaram-se os dados de Severidade média de cada um dos postos das sub-regiões para ajustá-los a funções distribuição de probabilidades. A Tabela 4.4 destaca os p-valores (em negrito) usando o método de estimação de máxima verossimilhança para o aceite da melhor função distribuição de probabilidades para modelar a Severidade de secas em cada uma das três sub-regiões.

Tabela 4.4 - P-valores dos ajustes da Severidade de secas às funções distribuição de probabilidades

Distribuições	Sub-região 1	Sub-região 1	Sub-região 3
Qui-quadrado	< 0,0001	0,5471	< 0,0001
Erlang	< 0,0001	< 0,0001	< 0,0001
Exponencial	< 0,0001	< 0,0001	< 0,0001
Gama (2)	0,8616	0,9117	0,0896
GEV	0,7919	0,9735	0,2755
Gumbel	< 0,0001	< 0,0001	< 0,0001
Log-normal	0,4485	0,4817	0,1442
Logística	0,8567	0,9484	0,4825
Normal	0,5306	0,9514	0,0146
Weibull (2)	0,4095	0,9712	0,0141
Weibull (3)	0,1993	0,9512	0,0541

Fonte: Elaborada pelo autor

A Figura 4.9 exibe o ajuste dos dados da Severidade de seca (Função empírica) à função distribuição de probabilidades (Função teórica) para a sub-região 1. Aqui, o melhor ajuste foi verificado com a distribuição de probabilidade Gama com 2 parâmetros (Gama (2)), ver Tabela 4.4. Adicionalmente, foi aplicado o teste de Kolmogorov-Smirnov (KS) para averiguar a adequação desse ajuste. O resultado indicou também boa adequação da função teórica aos dados observada.

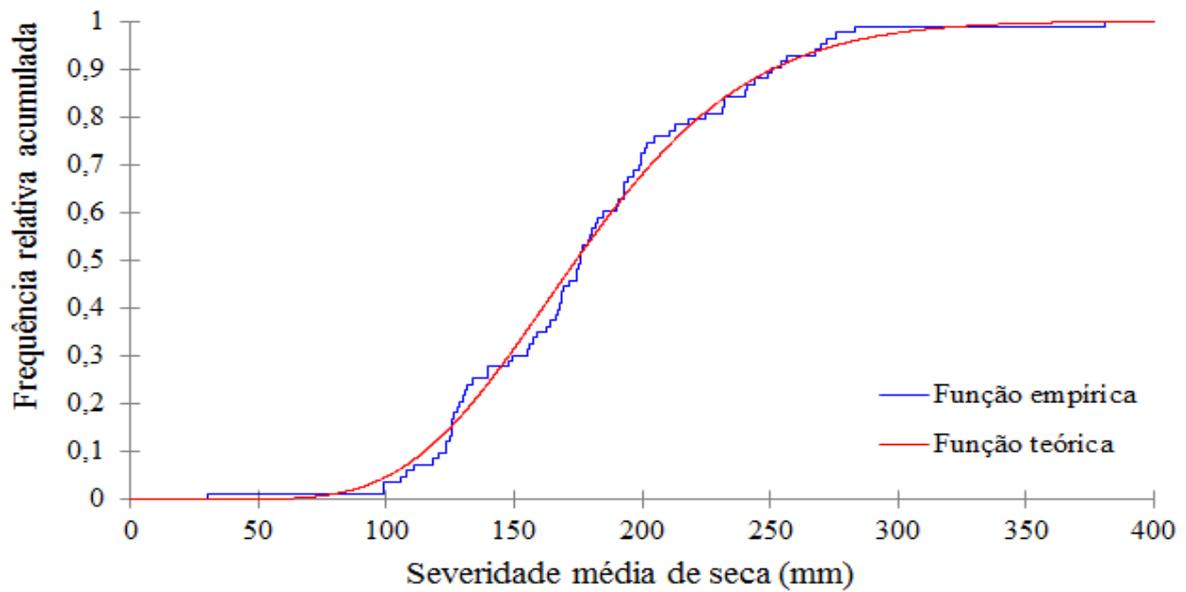


Figura 4.9: Função distribuição de probabilidades acumuladas Gama (2) para a sub-região 1

Na sub-região 2 a função distribuição de probabilidades que melhor se ajustou aos dados de severidade de seca foi a GEV (Generalized Extreme Value) ou Valor Extremo Generalizado, como pode ser visto pelo p-valor na Tabela 4.4. Essa função é bastante utilizada para o estudo de eventos de valores extremos. Como no primeiro caso, o teste de Kolmogorov-Smirnov (KS) também foi utilizado para averiguar a adequação do ajuste. O resultado indicou também boa adequação da função teórica aos dados observados. A Figura 4.10 exibe esse ajuste.

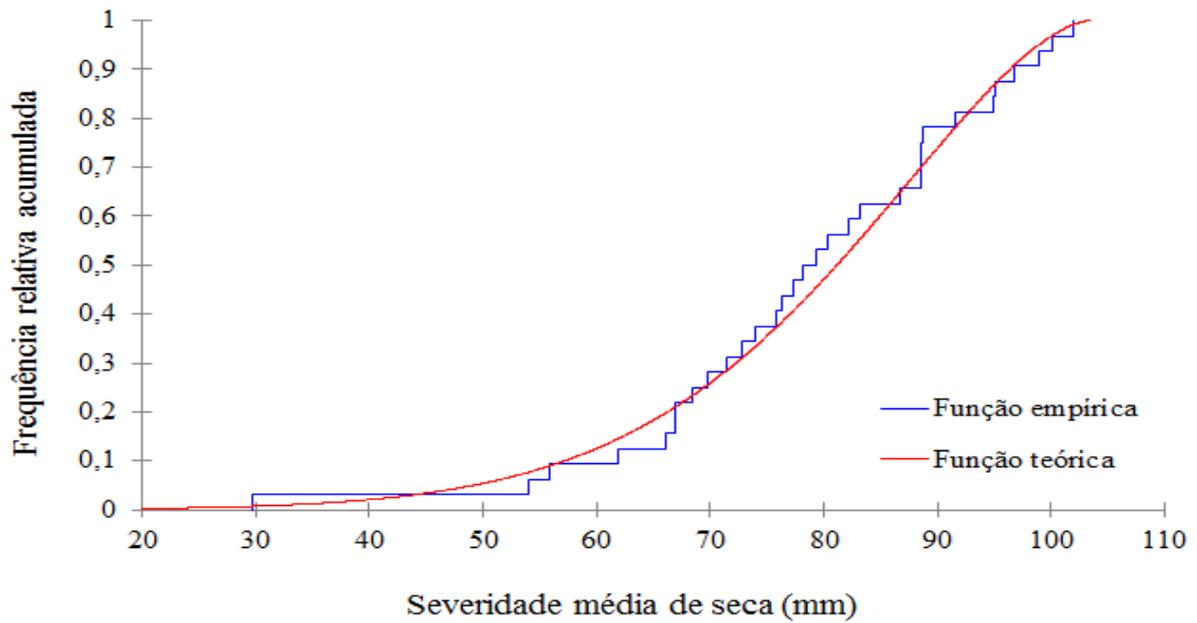


Figura 4.10: Função distribuição de probabilidades acumuladas GEV para a sub-região 2

Na sub-região 3 a função distribuição de probabilidades que melhor representa os dados da Severidade de secas (ver Tabela 4.4) é a Logística. Também, aqui o teste de Kolmogorov-Smirnov foi utilizado para testar a aderência da função empírica à função teórica. Essa aderência pode ser vista na Figura 4.11.

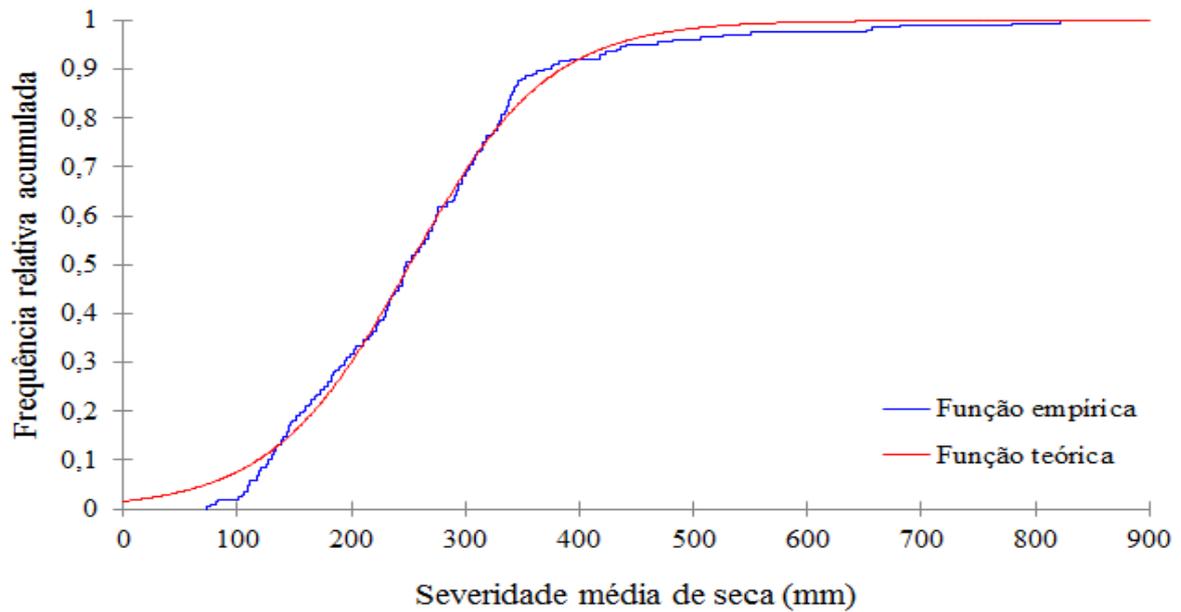


Figura 4.11: Função distribuição de probabilidades acumuladas Logística para a sub-região 3

Ajuste de funções distribuição de probabilidades marginal à variável Duração de secas

Após observar o comportamento da Duração de secas nas três sub-regiões, utilizaram-se os dados de Duração média de secas de cada um dos postos das sub-regiões para ajustá-los a funções distribuição de probabilidades. A Tabela 4.5 destaca os p-valores (em negrito) usando o método de estimação de máxima verossimilhança para o aceite da melhor função distribuição de probabilidades para modelar a Duração de secas em cada uma das três sub-regiões.

Tabela 7 - P-valores dos ajustes da Duração de secas às funções distribuição de probabilidades

Distribuições	Sub-região 1	Sub-região 1	Sub-região 3
Qui-quadrado	< 0,0001	< 0,0001	< 0,0001
Erlang	< 0,0001	< 0,0001	< 0,0001
Exponencial	< 0,0001	< 0,0001	< 0,0001
Gama (2)	0,4769	0,4090	0,7151
GEV	0,6847	< 0,0001	0,6497
Gumbel	< 0,0001	< 0,0001	< 0,0001
Log-normal	0,5807	0,7049	0,3841
Logística	0,4801	0,5221	0,4404
Normal	0,2222	0,2686	0,4473
Weibull (2)	0,2115	0,1163	0,1168
Weibull (3)	0,4701	0,9921	0,3739

Fonte: Elaborada pelo autor

A Figura 4.12 exibe o ajuste dos dados de Duração de secas (Função empírica) à função distribuição de probabilidades (Função teórica) para a sub-região 1. Neste caso, o melhor ajuste foi verificado com a distribuição de probabilidade GEV, ver Tabela 4.5. Adicionalmente foi aplicado o teste de Kolmogorov-Smirnov (KS) para averiguar a adequação desse ajuste. O resultado indicou também boa adequação da função teórica aos dados observados.

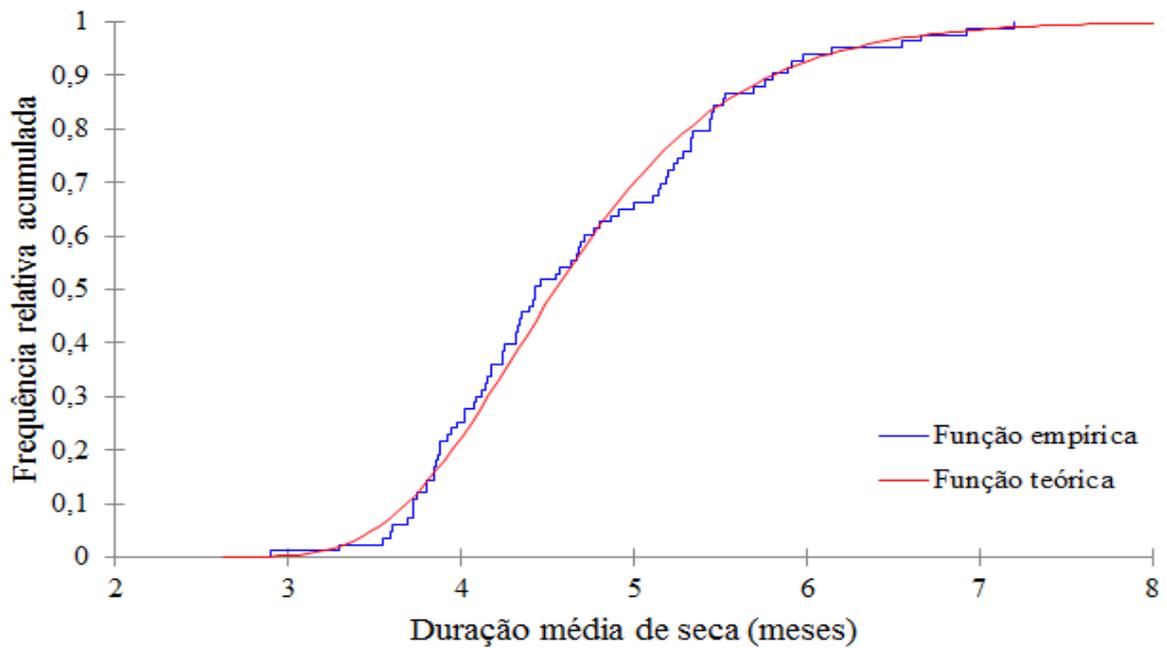


Figura 4.12: Função distribuição de probabilidades acumuladas GEV para a sub-região 1

Na sub-região 2 a função distribuição de probabilidades que apresentou melhor se ajustou aos dados de Duração de secas foi a Weibull com 3 parâmetros, como pode ser visto pelo p-valor na Tabela 4.5. Como no primeiro caso, o teste de Kolmogorov-Smirnov (KS) também foi utilizado para verificar a adequação do ajuste. O resultado indicou também boa adequação da função teórica aos dados observados. A Figura 4.13 exibe esse ajuste.

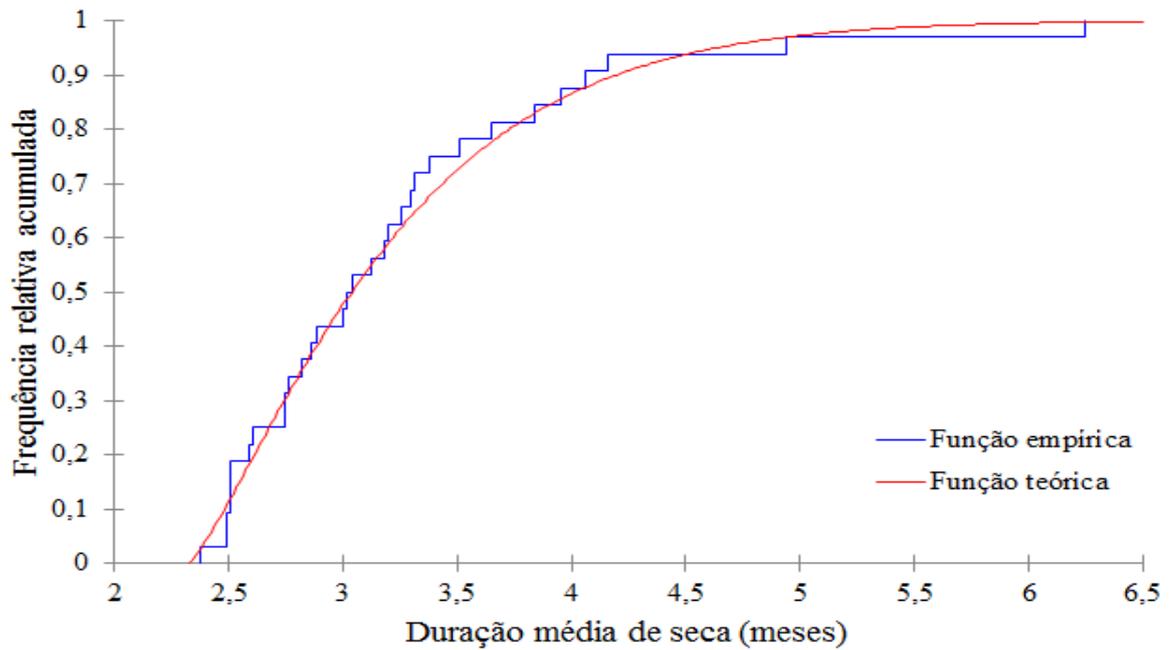


Figura 4.13: Função distribuição de probabilidades acumuladas Weibull (3) para a sub-região 2

Na sub-região 3 a função distribuição de probabilidades Gama com 2 parâmetros apresentou o ajuste mais satisfatório para modelar a Duração de secas que as demais (Tabela 4.5) A distribuição de Weibull é uma distribuição assimétrica e é bastante utilizada em estudos de índices pluviométricos (González e Donaire, 2000; Amorim e Pereira, 2008; Euclides et al., 2001). Também, aqui o teste de Kolmogorov-Smirnov foi utilizado para testar a aderência da função empírica à função teórica. Essa aderência pode ser vista na Figura 4.14.

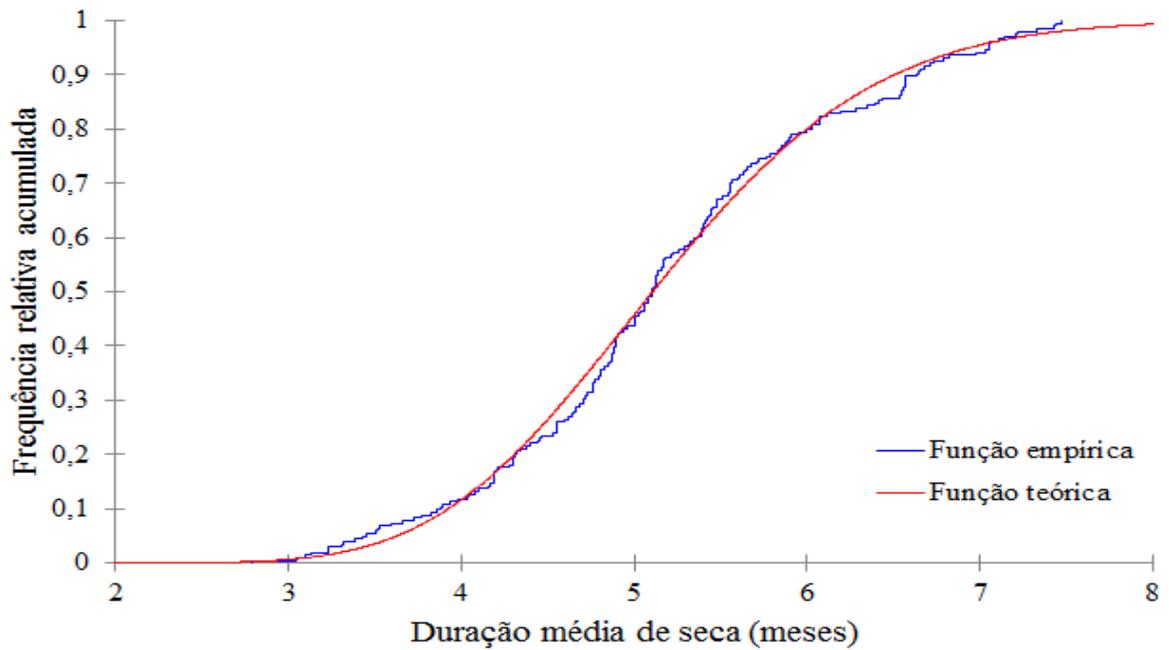
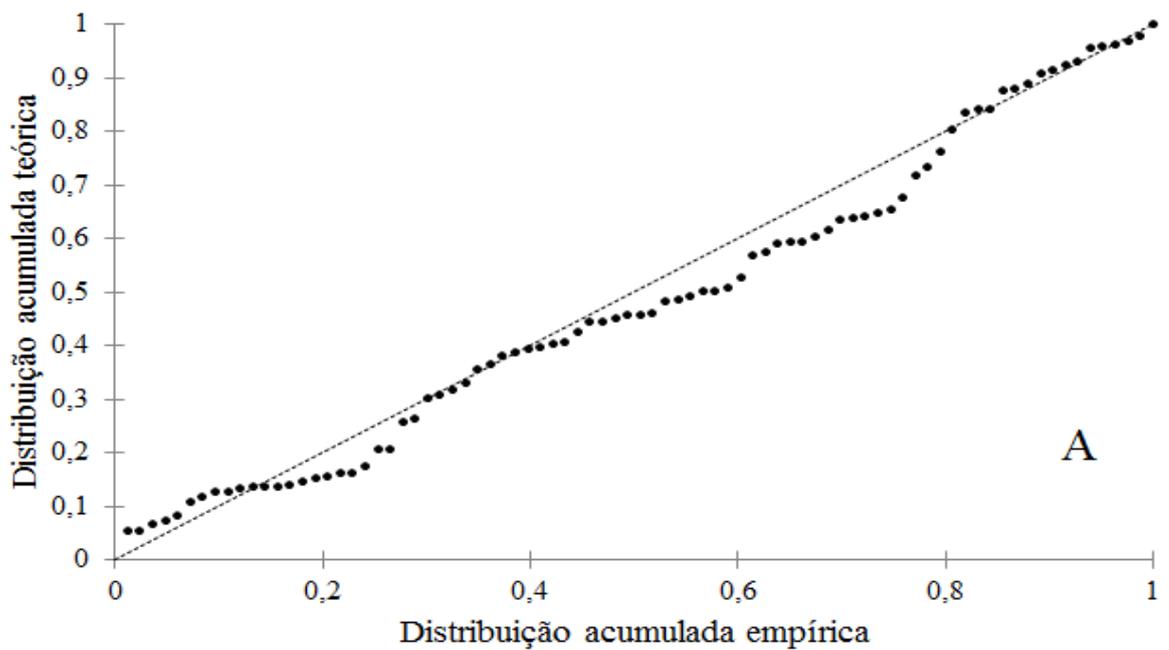


Figura 4.14: Função distribuição de probabilidades acumuladas Gama (2) para a sub-região 3

A Figura 4.15 compara a distribuição de probabilidades acumulada empírica e teórica da Severidade de seca na sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C). Destacam-se o bom ajuste para as sub-regiões 1 e 2 e um ótimo ajuste para a sub-região 3.



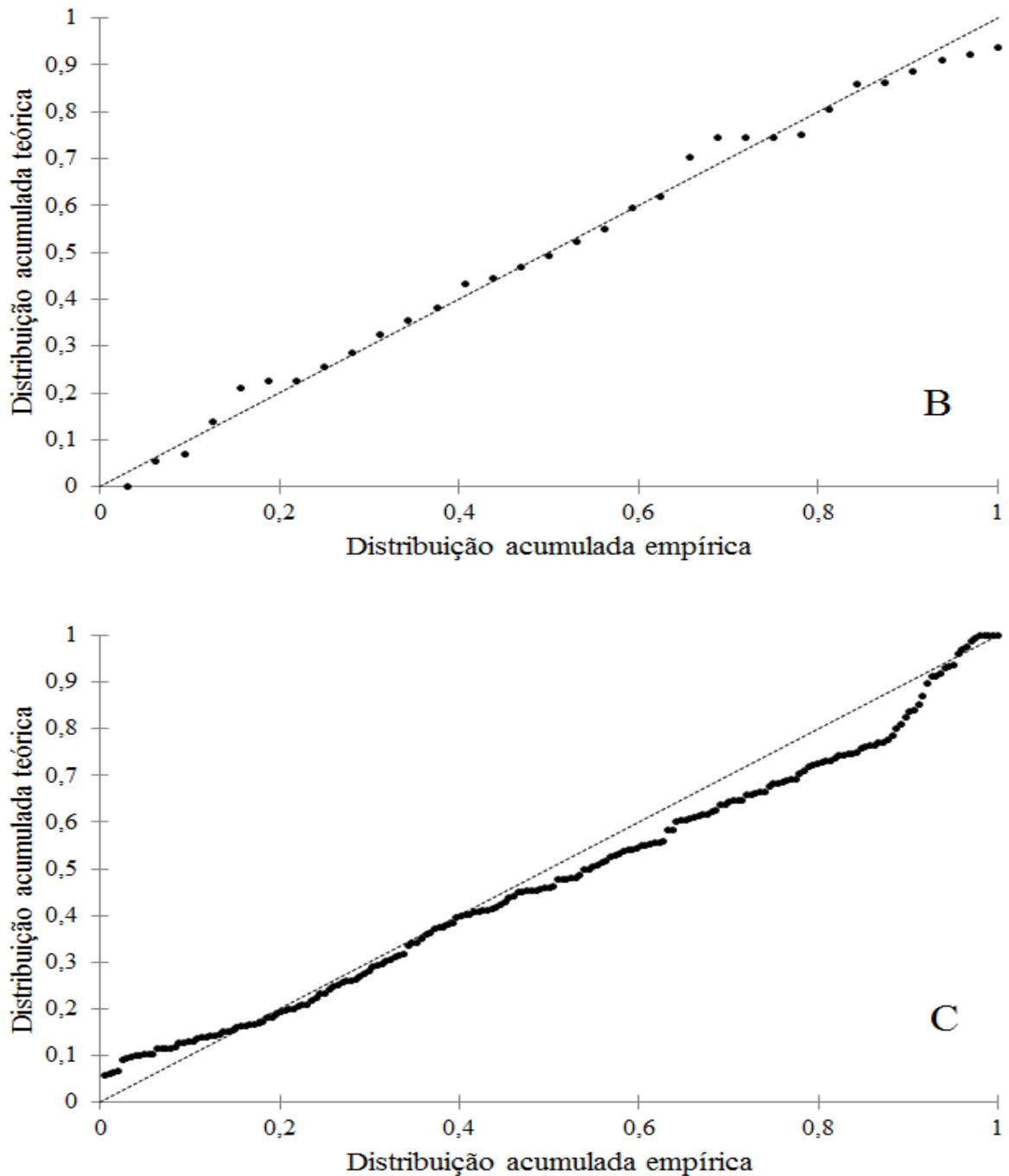
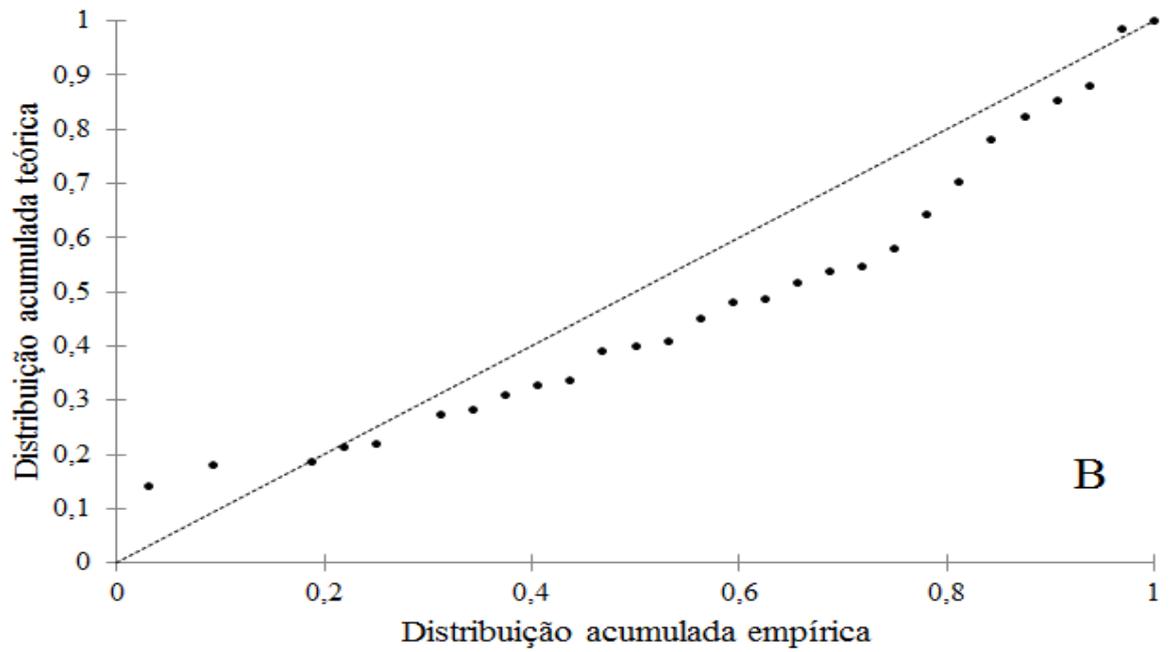
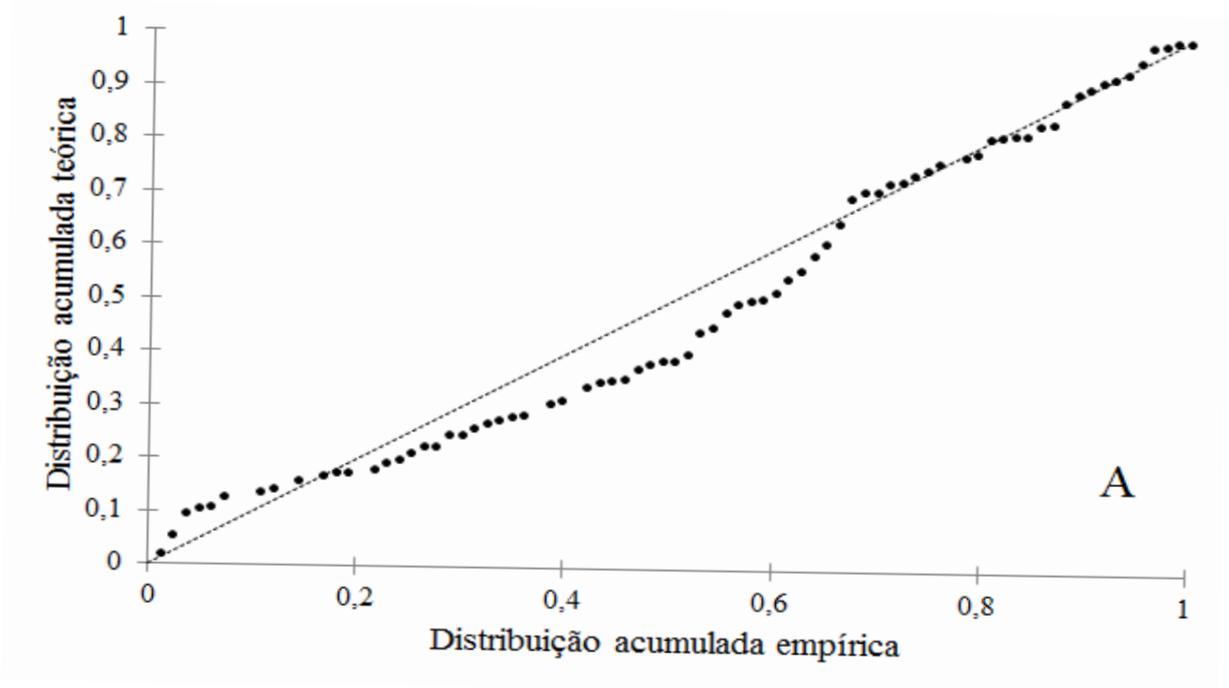


Figura 4.15: Distribuição de probabilidades empírica e teórica da Severidade de seca da sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C)

Já a Figura 4.16 compara a distribuição de probabilidades acumulada empírica e teórica da Duração de seca na sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C). Destacam-se o bom ajuste para as três sub-regiões.



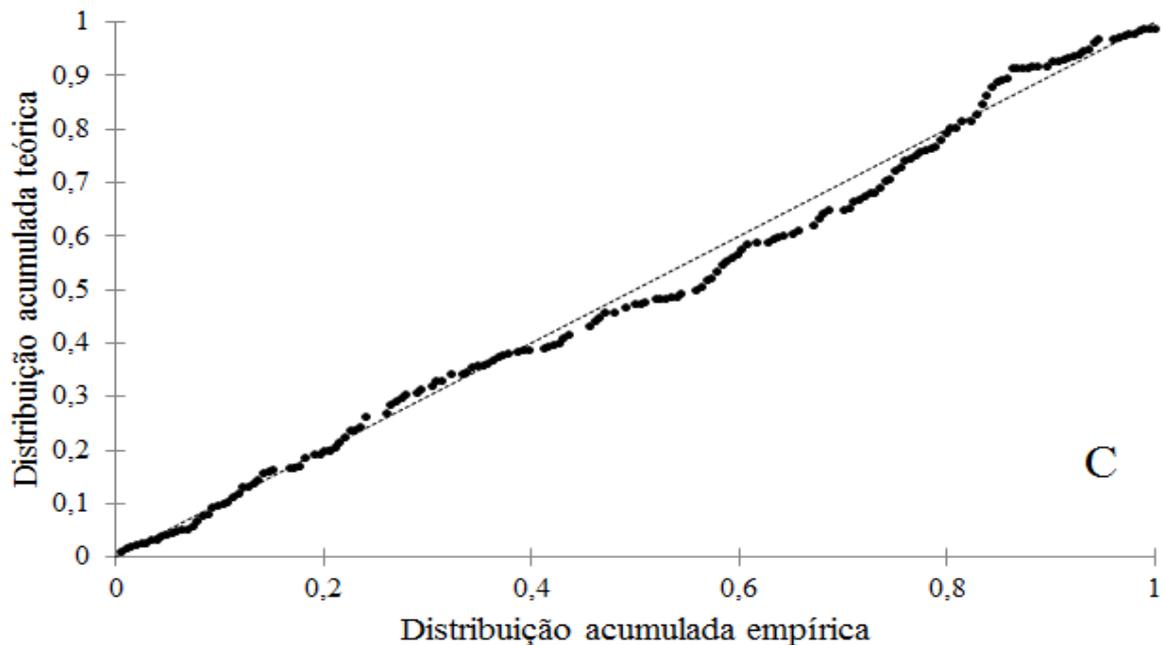


Figura 3: Distribuição de probabilidades empírica e teórica da Duração de seca da sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C)

Ajuste da distribuição de probabilidades Normal bivariada às variáveis Severidade e Duração de secas

Adicionalmente, foram utilizados os valores das variáveis Durações (meses) e Severidade (mm) de secas para ajustá-los a uma distribuição de probabilidades Normal bivariada. O interesse aqui é o de observar o comportamento do ajuste bivariado, embora superficialmente estudado. Inicialmente, para se utilizar a Normal bivariada é necessário estimar os coeficientes de correlação entre as variáveis Duração e Severidade de secas. Na Tabela 4.6 podem ser vistos os valores desses coeficientes de correlação.

Tabela 4.6 - Correlação entre os valores de Duração (meses) e Severidade (mm) e de secas para as três sub-regiões

Áreas	Correlação
Sub-região 1	0,61
Sub-região 2	0,24
Sub-região 3	0,77

Pode ser visto na Tabela 4.6 que a sub-região 3 apresenta o maior valor para o coeficiente de correlação entre as variáveis e a sub-região 2 o mais baixo deles. Em geral, esses valores são muito baixos para que haja a hipótese de ajustamento das variáveis à Normal bivariada.

As Figuras 4.17, 4.18 e 4.19 apresentam a relação espacial entre os valores das variáveis Duração (meses) e Severidade (mm) de secas para o período de 1984 a 2014, para as sub-regiões 1, 2 e 3, respectivamente. Na sub-região 1 a Duração de secas é sempre superior à Severidade de secas.

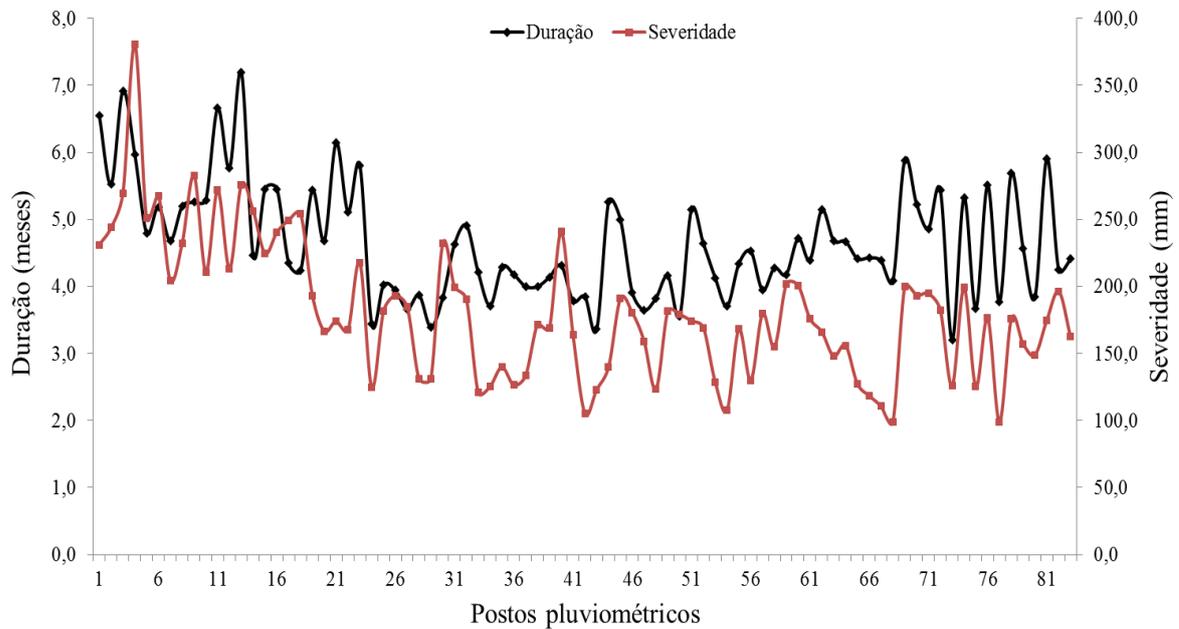


Figura 4.17: Ciclo da variabilidade e da duração e severidade da sub-região 1

A distribuição da duração e a severidade da seca para o período em estudo na sub-região 2 está representada na figura 4.18 composta de 32 postos pluviométricos. Destaca-se que as variabilidades das durações e severidades são proporcionais entre os postos 15 a 20, e as irregularidades com maiores amplitudes ocorreram nas maiorias dos postos estando relacionadas aos fenômenos de meso, micro e grandes escalas atuantes na área.

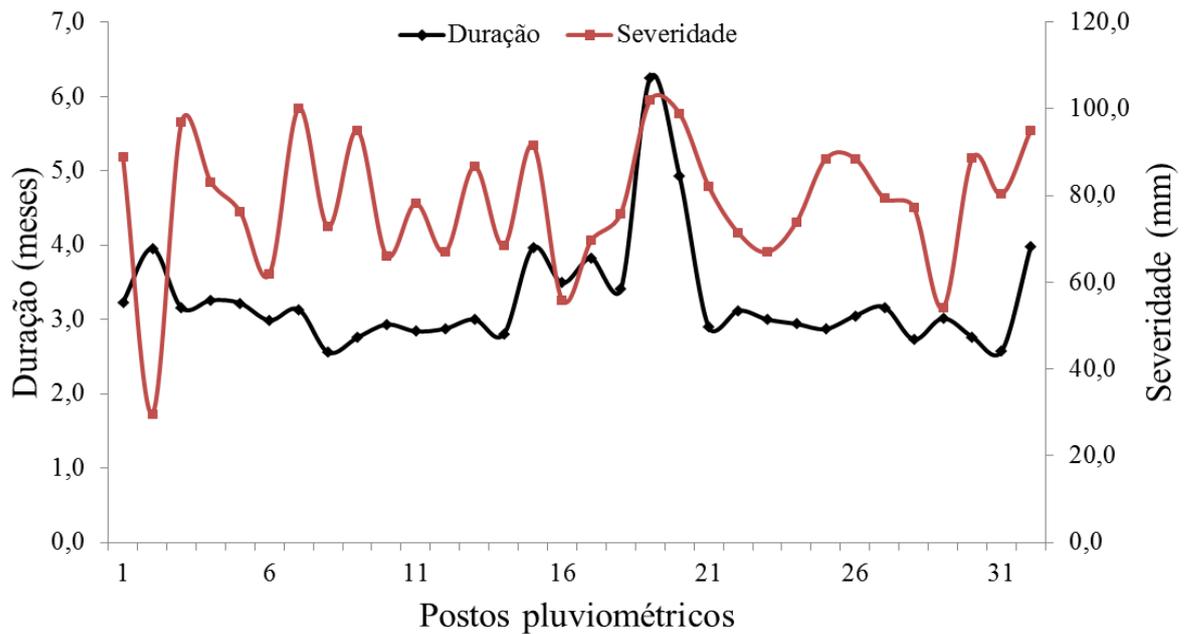


Figura 4: Ciclo da variabilidade e da duração e severidade da sub-região 2

A Figura 4.19 representa a distribuição da duração e a severidade da seca para o período em estudo na sub-região 3, composta de 204 postos pluviométricos. As durações e as severidades apresentaram-se com altas irregularidades. Essas irregularidades se devem a ocorrência de longos períodos com seca motivada pela atuação do El Niño e dos fatores de média e pequena escalas. Destacam-se os postos 6 a 11 e o posto 81 onde suas durações e severidades foram mais regularizadas quando comparados aos demais postos em estudo.

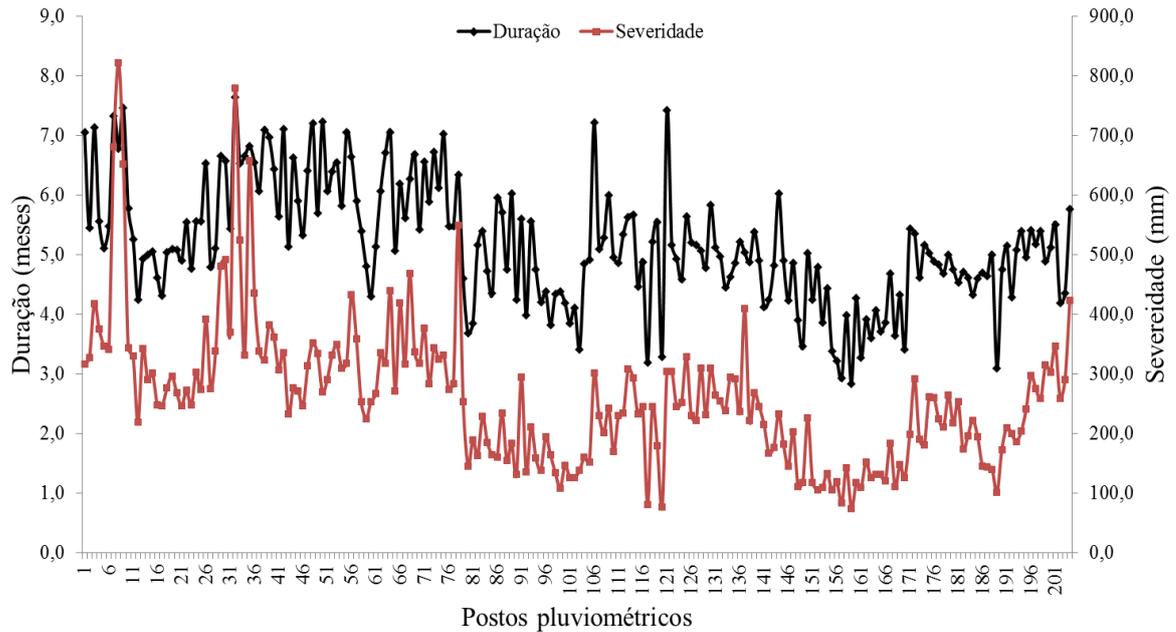
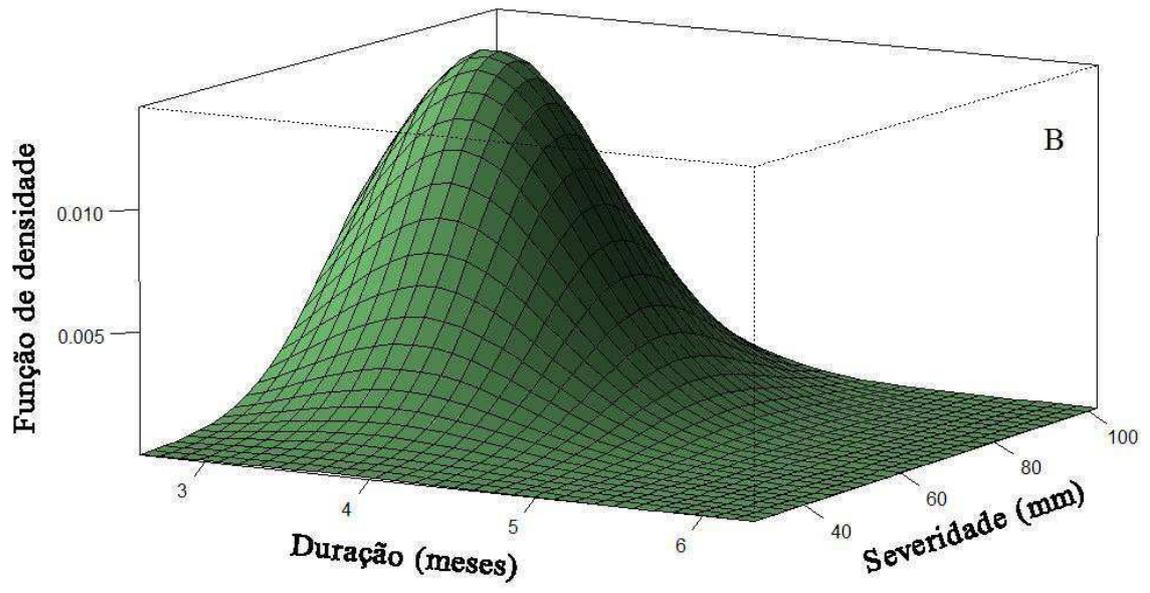
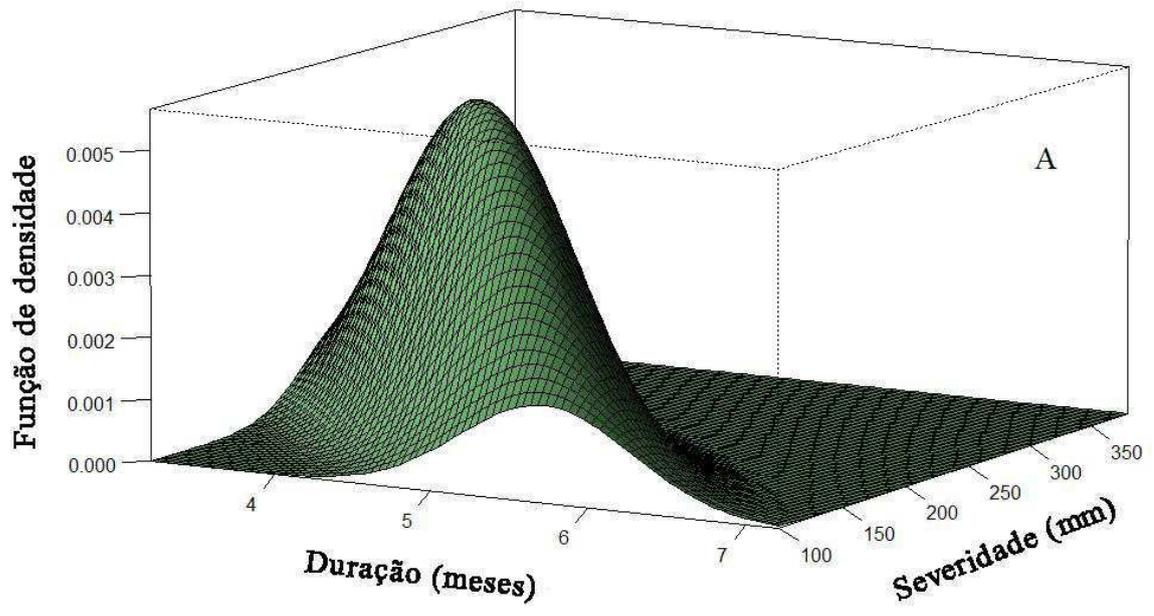


Figura 5.19: Ciclo da variabilidade e da duração e severidade da sub-região 3

Para cada uma das três regiões, ajustou-se uma distribuição normal bivariada, com o objetivo de verificar o comportamento das mesmas, pois em um primeiro teste observou-se que as variáveis de duração e severidade de seca apresentam uma boa correlação.

A Figura 4.20 apresenta o ajuste dos dados de Severidade e Duração de secas para a sub-região 1 (A), sub-região 2 (B), sub-região 3 (C) à função distribuição de probabilidades Normal bivariada.

Essas probabilidades de superfície se relacionam com a distribuição bivariada. Ela também informa como duas variáveis são correlacionadas uma a outra. O gráfico da Normal bivariada é composto por três dimensões: duas dimensões se referem as variáveis, neste caso, Severidade e Duração de secas e a terceira dimensão representa a função densidade de probabilidades Normal bivariada.



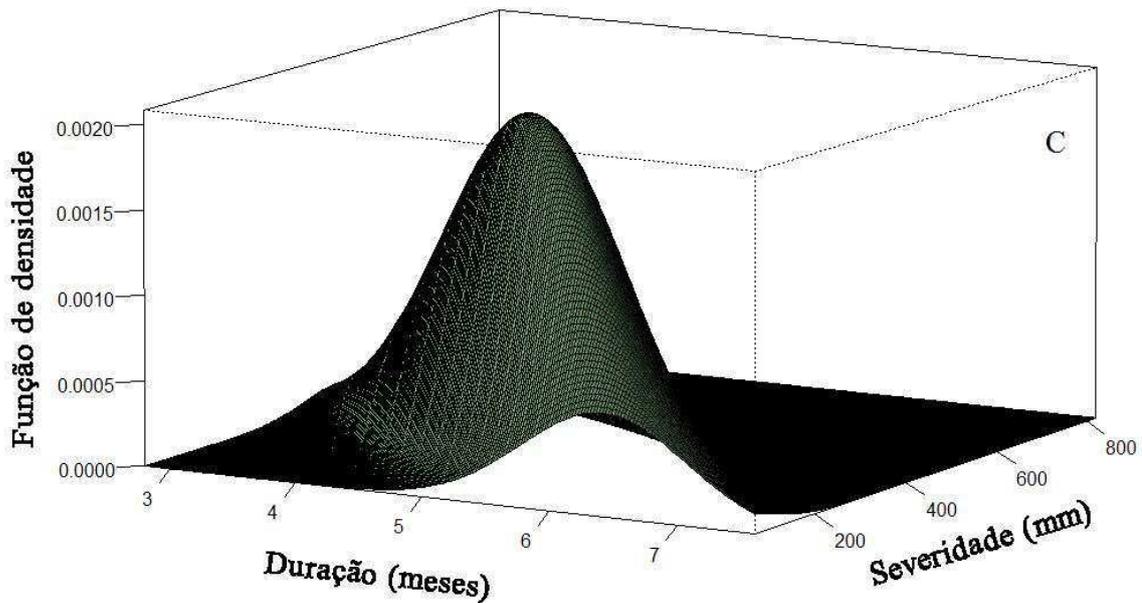


Figura 6: Ajuste da distribuição Normal bivariada: (A) sub-região 1 (B) sub-região 2 e (C) sub-região 3

A Tabela 4.7 exibe os valores do teste de Henze-Zirkler (HZ). O objetivo deste teste é verificar a normalidade dos parâmetros das amostras das variáveis em estudo e se essas amostras seguem uma distribuição Normal bivariada ao nível de significância 0,05. Haja vista que os p valores são significativamente inferiores a 0,05 pode-se concluir que as amostras de dados não se ajustam à Normal bivariada.

Tabela 4.7 - Teste MVN de Henze-Zirkler

Área do teste	Estatística HZ	P-valor
Sub-região 1	1,12	0,01
Sub-região 2	1,98	$3,9 \times 10^{-5}$
Sub-região 3	3,41	$4,8 \times 10^{-8}$

Adicionalmente ao teste estatístico de Henze-Zirkler (HZ), foram plotados os valores das variáveis Severidade e Duração de secas em gráficos denominados Qui-quadrado (Figura 4.21). Nesse tipo de gráfico quando os pontos se alinham segundo uma reta a suposição de normalidade é sustentada. Caso contrário, quando os pontos se afastam da reta sugere-se que há falta de normalidade. Assim, foi possível afirmar que os dados de Duração e Severidade de secas das três sub-regiões estudadas não seguem uma distribuição normal bivariada.

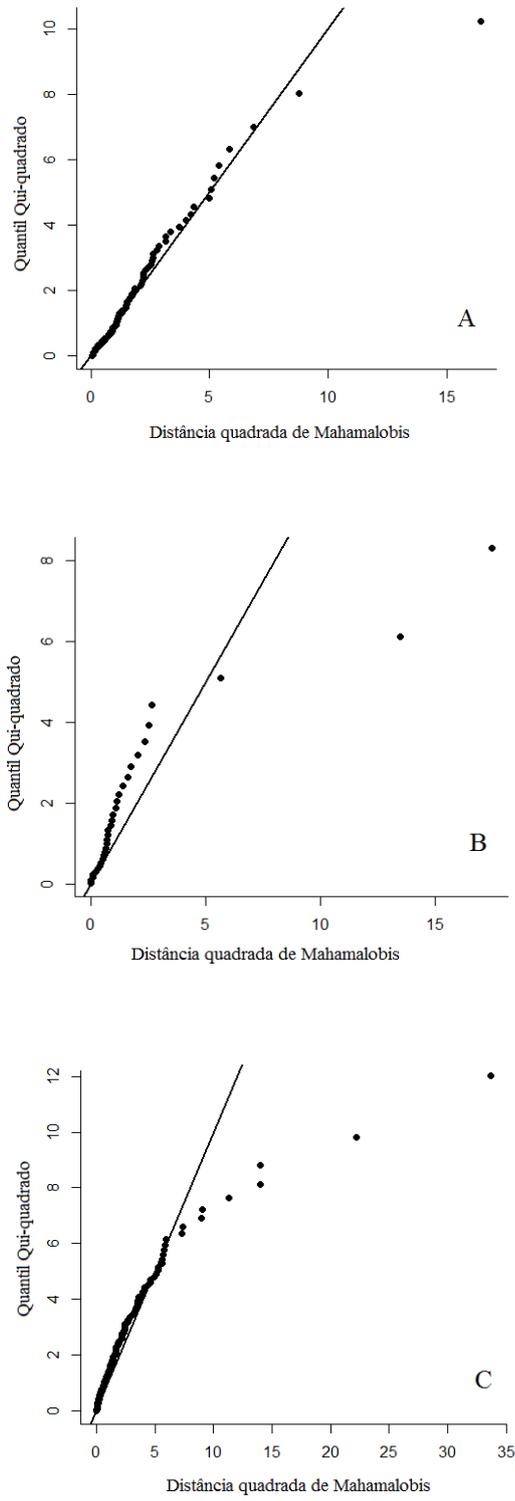


Figura 7: Q-Q Plot bivariado para a sub-região 1 (A), sub-região 2 (B) e sub-região 3 (C)

5 CONCLUSÕES

Os padrões das sub-regiões pluviometricamente homogêneas obtidos com base na Análise de Agrupamento exibem três sub-regiões com características de precipitação distintas para o trimestre mais chuvoso. Com relação aos períodos secos, destacam-se os cinco meses da sub-região 3, de maio a setembro. O período mais chuvoso desta sub-região 3 também difere das demais. Isso ocorre porque essa sub-região está localizada no sul do SAB e é influenciada pelas Frentes Frias oriundas do Sul do Brasil. Já os períodos chuvosos das sub-regiões 1 e 2 estão relacionados aos sistemas de grandes escalas como a Zona de Convergência Intertropical (ZCIT) e os Distúrbios Ondulatórios de Leste (DOL).

Com base nas variáveis Severidade e Duração de secas e na Análise de Agrupamento foram delimitadas três sub-regiões homogêneas para cada uma das variáveis. Como visto nas Figuras 4.6 e 4.8, respectivamente, as sub-regiões de Severidade das secas não guardam muitas semelhanças com as sub-regiões de Duração das secas. Isso significa que podem ocorrer secas muito severas de curta duração ou secas menos severas de longa duração.

Os ajustes das funções distribuição de probabilidades univariadas aos valores das variáveis Severidade e Duração das secas para cada uma das sub-regiões podem ser considerados satisfatórios. Já os ajustes das variáveis Severidade e Duração das secas à distribuição de probabilidades Normal bivariada para as três sub-regiões não lograram êxito, fato que já era esperado.

As funções distribuição de probabilidades Gama, GEV e Logística foram os modelos que melhor descrevem a Severidade de secas nas sub-regiões 1, 2 e 3, respectivamente. E as funções distribuição de probabilidades GEV, Weibull três parâmetros e Gama dois parâmetros, foram os modelos que melhor representaram a Duração de secas nas sub-regiões 1, 2 e 3, respectivamente.

6 SUGESTÕES PARA TRABALHOS FUTUROS

A sugestão para trabalho futuro é o ajustamento de funções distribuição de probabilidades marginais para as variáveis Severidade e Duração de secas e utilizar uma função cópula para obter o modelo bivariado capaz de representar a variabilidade conjunta dessas duas variáveis.

7 REFERÊNCIAS

- AGÊNCIA NACIONAL DE ÁGUAS (ANA). Sistema de informações hidrológicas (HidroWeb). Disponível em: < <http://hidroweb.ana.gov.br/>>. Acesso em 26 abr. 2013.
- AGRAWALA, S.; BROAD, K.; GUSTON, D. H. *Integrating climate forecasts and societal decision making: Challenges to an emergent boundary organization*. Science Technology Human Values, 26: p.454 – 477, 2001.
- ALLEY, W. M. *The palmer drought severity index: limitations and assumptions*. Journal of Climate and Applied Meteorology, Boston, v. 23, n. 7, p. 1100-1109, 1984.
- AMORIM, S.V. de, PEREIRA, D.J. de A. *Estudo comparativo dos métodos de dimensionamento para reservatórios utilizados em aproveitamento de água pluvial*. Revista Ambiente Construído 8, 53-66.
- ASHRAF, M.; ROUTRAY, J. K. *Perception and understanding of drought and coping strategies of farming households in north-west Balochistan*. International Journal of Disaster Risk Reduction, p.14 - 30, 2013.
- ASSIS, F. N. DE., ARRUDA, H.V., PEREIRA, A. R. *Aplicações estatísticas a climatologia*. Pelotas, Ed. Universitária/UFPel, 1996, 161p.
- BHALME, H. N. MOOLEY, D. A. *Large scale drought/floods and monsoon circulation*. Mon. Wea. Rev., v.8, p.1197-1211. 1980.
- BRUCE, J. P. *Natural disaster reduction and global change*. Bulletin of the American Meteorological Society, n.75, p.1831–1835. 1994.
- BUSSAB W. O.; MIASAKI, E. S. e ANDRADE, D. F. *Introdução à análise de agrupamentos*. In: Simpósio Nacional de Probabilidade e Estatística, 9. São Paulo, ABE. 105p. 1990.
- BYUN, H. R.; WILHITE, D. A. *Objective quantification of drought severity and duration*. Journal of Climate, Boston, v. 12, n. 9, p. 2747-2756, Sept. 1999.
- CAMARGO, M. B. P., HUBBARD, K. G. *Spatial and temporal variability of daily weather variables in sub-humid and semi-arid areas of the U.S. High Plains*. Agricultural and Forest Meteorology., 93, 141-148. 1999.
- CANUTO, A.; LUZ, C. R. S.; LAZZARIN, F. *Conflitos no campo – Brasil 2012*. Goiânia-GO: CPT Nacional. 2013. 188p.
- CASTRO, J. Geografia da fome. 9 editora. Rio de Janeiro: Civilização Brasileira, 2008. 318p.

CHANGNON, S. A.; PIELKE JR, R. A.; CHANGNON, D.; SYLVES, R. T.; PULWARTY, R. *Human factors explain the increased losses from weather and climate extremes*. Bulletin of the American Meteorological Society, 81 (3), p.437 – 442, 2000.

CHAVES, R. R.; CAVALCANTI, I. F. A. *Atmospheric Circulation Features Associated with Rainfall Variability over Southern Northeast Brazil*. Monthly Weather Review. v.129, p. 2614-2626, 2001.

CHOW, V. T.; MAIDMENT, D. R.; Mays, L. W. *Applied Hydrology*. McGraw-Hill Book Company, New York, 1998.

CLAUSE, B.; PEARSON, C. P. *Regional frequency analysis of annual maximum streamow drought*. Journal of Hydrology, 173, p.111 – 130. 1995.

COELHO, C. A. S.; UVO, C. B.; AMBRIZZI, T. *Exploring the impacts of the tropical Pacific SST on the precipitation patterns over South America during ENSO periods*. Theor. Appl. Climatol. v. 71, p. 185-197, 2002.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). Boletim de Grãos. Setembro de 2012. Disponível em: <<http://www.conab.gov.br>. Acesso em 20 março de 2015.

COOK, E. R.; SEAGER, R.; CANE, M. A.; STAHL, D. W. *North American drought: reconstructions, causes, and consequences*. Earth Science Reviews, v.81, p.93 – 134. 2007.

COUTO, H. T. Z. *Distribuições de diâmetro em plantações de pinus caribaea morelet*. Tese de Mestrado, ESALQ, 1980, 79p.

DRACUP, J. A.; LEE, K. S.; PAULSON, E. G. *On the statistical characteristics of drought events*. Water Resources, 16 (2), p.289–296, 1980.

DUARTE, R. S. *Do desastre natural à calamidade pública: a seca de 1998-1999*. Fortaleza: Banco do Nordeste. Recife: Fundação Joaquim Nabuco, 2002. 144p.

EDWARDS, D. C.; MCKEE, T. B. *Characteristics of twentieth century drought in the United States at multiple time scales*. Fort Collins: Colorado State University, 1997. (Climatology Report, n. 97–2).

EISCHEID, J. K., BAKER, C. B., KARL, T., DIAZ, H. F. *The quality control of long-term climatological data using objective data analysis*. Journal of Applied Meteorology, 34, 2787-2795. 1995.

UCLYDES, H.P., FERREIRA, P.A., RUBERT, O.A.V., SANTOS, R.M. dos, 2001. *Regionalização hidrológica na bacia do alto São Francisco a montante da barragem de Três Marias- Minas Gerais*. Rev. Bras. de Recursos Hídricos 6, 81-105.

EUROPEAN COMMUNITIES. *Addressing the Challenge of Water Scarcity and Droughts in the European Union*. Commun. Com. 414 Final, Brussels. 2007.

- EVERITT, B.S.; DUNN, G. *Applied multivariate analysis*. London: Edward Arnold, 1991. 400p.
- FAO. Report of FAO-CRIDA. *Expert Group Consultation on Farming System and Best Practices for Drought-prone Areas of Asia and the Pacific Region*. Food and Agricultural Organization of United Nations. Published by Central Research Institute for Dryland Agriculture, Hyderabad, India. 2002.
- FEYEN, L.; DANKERS, R. *Impact of global warming on streamflow drought in Europe*. Journal of Geophysical Research Atmospheres. v.114, D1716, p.78 – 92, 2009.
- FERREIRA, A. G.; MELLO, N. G. S. *Principais sistemas atmosféricos atuantes sobre a região Nordeste do Brasil e a influência dos oceanos Pacífico e Atlântico no clima da região*. Revista Brasileira de Climatologia, v.1, n.1. p.15 - 28. 2005.
- FREITAS, M. A. S. *Um sistema de suporte à decisão para o monitoramento de secas meteorológicas em regiões semiáridas*. Revista Tecnologia, Fortaleza, v. Suplemento, p.84 - 95, 2005.
- GANDIN, L. S. *Complex quality control of meteorological observations*. Monthly Weather Review, 116, 1137-1156. 1998.
- GOIS, G. de. *Análise espacial de eventos de secas com base no índice padronizado de precipitação e análise de agrupamento*. Revista Brasileira de Ciências Ambientais. n.23, p.55 - 65. 2012.
- GONZÁLEZ, M.E.P.E., DONAIRE, J.J.S. *Distribuciones estadísticas ajustadas a las series temporales de totales anuales de precipitación en España: aspectos geográficos*. Rev. Geographicalia 38, 11-29.
- GREENLAND, S. and FINKLE, W.D. *A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses*. American Journal of Epidemiology, 142, 1255-1264. 1995.
- GUTTMAN, N. V. Quayle, R. G. 1990. *A review of cooperative temperature data validation*. Journal of Atmospheric and Oceanic Technology., 7, 334-339.
- HAAN, C. T. *Statistical methods in hidrology*. Ames, Iowa State University Press, 1977, 378p, il.
- HASTENRATH, S. *Explorin the climate problems of Brazil's Nordeste: a review*. Climatic Change, p. 243-251, 2012.
- HASTINGS, N.A.J.; PEACOCK, J.B. *Statistical distributions: a handbook for students and practitioners*. New York: J. Wiley, 1975.
- HAYES, M. J.; SVOBODA, M.D.; WILHITE, D.A.; VANYARKHO, O.V. *Monitoring the 1996 drought using the Standardized Precipitation Index*. Bulletin of the American Meteorology Society, v. 80, n. 3, p.429-438, 1999.

- HEIM JUNIOR, R. R. H. *A review of twentieth - century drought indices used in the United States*. Bulletin of the American Meteorology Society, v. 83, n.8, p.1149-1163, 2002.
- HOSKING, J. R. M.; WALLIS, J. R. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, New York. 1997.
- HUBBARD, K. G., GODDARD, S., SORENSEN, W. D. Wells, N., Osugi, T. T. 2005. *Performance of Quality Assurance Procedure for an Applied Climate Information System*. Journal of Atmospheric and Oceanic Technology., 22, 105-112.
- 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.
- JENKINSON, A. F. *The frequency distribution of the annual maximum (or minimum) of meteorological elements*. Quarterly Journal of the Royal Meteorological Society, London, v.81, n.551.501.45, p.158-171, 1955.
- JOHNSON, N. L e KOTZ, S. *Distribution in statistics, continuo univariate distribution*. New York, Houghton Mifflin, 1970, 2v.
- KANE, R. P. *Prediction of droughts in North-east Brazil: Role of ENSO and use of periodicities*. Internacional Journal of Climatology, v.17, p.655-665,1997.
- KAYANO, M. T; ANDREOLI R. V. *Relationships between rainfall anomalies over northeastern Brazil and El Niño-Southern Oscillation*. Jornal of Geophysical Research, v.111, D13102, doi: 10.1029/2005JD006142, 2006.
- KAO, S.; GOVINDARAJU, R.S.; *A bivariate frequency analysis of extreme rainfall with implications for design*. Journal of Geophysical Research, n.112, D13110. doi:10.1029/2007JD008522. 2007.
- KIM, T. W.; VALDÉS, J. B.; YOO, C. *Nonparametric approach for estimating return periods of droughts in arid regions*. Journal of Hydrologic Engineering - ASCE 8, (5), p.237–246. 2003.
- KIM, T.-W.; VALDÉS, J. B.; YOO, C. *Nonparametric approach for bivariate drought characterization using Palmer drought index*. Journal Hydrologic Engineering, ASCE 11 (2), 134–143, 2006.
- KIM, D. H.; YOO, C.; KIM, T.-W. *Application of spatial EOF and multivariate time series model for evaluating agricultural drought vulnerability in Korea*. Advanced Water Resources, 34 (3), p.340 – 350, 2011.
- KITE, G. W *Frequency and risk analisys in hidrology*. Fort Colins, Water Resources Publications, 1978, v3, 395p.
- ORKMAZ, S. GOKSULUK, D. and ZARARSIZ, G. *MVN: Multivariate Normality Tests, 2015*. R package version 4.0.

- KOUSKY, V. E. *Frontal influences on northeast Brazil*. Monthly Weather Review, v.107, n. 9, p.1140 - 1153, 1979.
- KOUSKY, V. E.; CHU, P. S. *Fluctuations in annual rainfall for northeast Brazil*. J. Meteor. Soc. Japan, v. 57, p. 457-465, 1978.
- JONHSON, R. A. and WICHERN, D. W. *Applied Multivariate Statistical Analysis*. Sixth Ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2007.
- LE COMTE, D. *Weather highlights around the world*. Weatherwise. v.48, p.20 – 22, 1995.
- LUCIO, P. S.; CONDE, F. C.; CAVALCANTI, I. F. *Spatiotemporal monthly rainfall reconstruction via Artificial Neural Network*. Case study: south of Brazil. Advances in Costa, R. L.; Silva, F. D. S.; Sarmanho, G. F.; Lúcio, P. S. 675 Geosciences, v. 10, p. 67-76. 2007.
- MARENGO, J. A.; ALVES, L. M.; BESERRA, E. A.; LACERDA, F. F. *Variabilidade e mudanças climáticas no semiárido brasileiro*. In: MEDEIROS, S. S.; GHEYI, H. R.; GALVÃO, C. O.; PAZ, V. P. S. (orgs). Recursos hídricos em regiões áridas e semiáridas. Campina Grande, PB: Instituto Nacional do Semiárido, 2011. 440p.
- MCKEE, T. B.; DOESKEN, N. J.; KLEIST, J. *The relationship of drought frequency and duration of time scales*. Eighth Conference on Applied Climatology. American Meteorological Society. Jan. 17-22, Anaheim, CA, p.179-186, 1993.
- MEEK, D. W. Hatfield, J. L. 1994. *Data quality checking for single station meteorological databases*. Agricultural and Forest Meteorology., 69, 85-109.
- MELO, R. O. *Caracterização da Produção de Sedimentos de uma Microbacia no Semiárido Pernambucano*. Dissertação (Mestrado em Ciência do Solo). Universidade Federal Rural de Pernambuco. 2008, 94p.
- MELO FILHO, J. F.; SOUZA A. L. V. *O manejo e a conservação do solo no Semiárido baiano: desafios para a sustentabilidade*. Bahia Agrícola. v.7, n.3, 11p, 2006.
- MISHRA, A. K.; SINGH, V. P. *A review of drought concepts*. Journal of Hydrology, v. 391. p. 202–216. 2010.
- MIRSHAWKA, V. *Estatística*. Vol II, São Paulo, Liv. Novel, 1971, 367p.
- MOLION, L. C. B.; BERNARDO, S. O. *Uma revisão da dinâmica das chuvas no Nordeste brasileiro*. Revista Brasileira de Meteorologia, v.17, n.1, p.1-10, 2002.
- MORRISON, D. F. *Multivariate statistical methods*. 3a ed. McGraw-Hill Publishing. 459p. 1990.
- MOSCATI, M. C. L.; GAN, M. A. *Rainfall variability in the rainy season of semiarid zone of Northeast Brazil (NEB) and its relation to wind regime*. Int. J. Climatol. 27: 493 – 512, 2007.

MOURA, A. D.; SHUKLA, J. *On the dynamics of droughts in Northeast Brazil: observations, theory and numerical experiments with a general circulation model*. J. Atmos. Sci., v. 38, p. 2653-2675, 1981.

NAE - Núcleo de Assuntos Estratégicos da Presidência da República. Mudança de Clima: Negociações internacionais sobre a mudança de clima; vulnerabilidade, impactos e adaptação à mudança de clima. Cadernos NAE, NAE-SECOM, Brasília, 2005, 250 p.

NANDINTSETSEG, B.; SHINODA, M. *Assessment of drought frequency, duration, and severity and its impact on pasture production in Mongolia*. Natural Hazards, 66: p.995 – 1008. 2013.

NOBRE, P.; MELO, A. B. C. *TVariabilidade Climática Intrasazonal Sobre o Nordeste do Brasil em 1998-2000*. Revista Climanálise, 2001. Cachoeira Paulista, SP. Disponível em: <<http://www6.cptec.inpe.br/revclima/revista/pdf/artigovariabilidadedez01.pdf>> :Acesso em : 20de junhode2006:

NOBRE, C. A.; MOLION, L. C. B. *The climatology of droughts and drought prediction, the effect of climatic variation on agriculture*. In: PARRY, M. P.; CARTER, T. R.; KONIJN, N. T. (Ed.). *Assessments in semiarid regions*, Dordrecht: Kluwer Academic Publishing, v.2, 1988.

NCDC (National Climatic Data Center) 2002. *US National Percent Area Severely to Extremely Dry and Severely to ExtremelyWet*. Disponível em: <http://www.ncdc.noaa.gov/oa/climate/research/txt>>. Acesso em: 30/07/2013.

OBASI, G. O. P. *WMO's Role in the International Decade for Natural Disaster Reduction*. Bulletin American Meteorological Society, n.75, p.1655 – 1661, 1994.

OLADIPO, E. O. *A comparative performance analysis of threemeteorological drought indices*. Journal of Climatology, Chichester, v. 5, n. 6, p. 655–664, Nov./Dec. 1985.

OLIVEIRA, A. S. de. *Interações entre sistemas frontais na América do Sul e a convecção da Amazônia*. 115f. Dissertação (Mestrado em Meteorologia), Instituto Nacional de Pesquisas Espaciais, São José dos Campos, INPE, 1986.

PALMER, W. C. *Keeping track of crop moisture conditions, nationwide: the new crop moisture index*. Weatherwise, 21, p. 156–161. 1968.

PIRES, V. *Frequência e intensidade de fenômenos meteorológicos extremos associados a precipitação*. Dissertação (Mestrado em Ciências e Engenharia da Terra) – Universidade de Lisboa, Portugal, 2003.

REBOITA, M. S.; GAN M. A.; ROCHA, R. P.; AMBRIZZI T. *Regimes de precipitação na América do Sul: Uma revisão bibliográfica*. Revista Brasileira de Meteorologia. v. 25, n. 2, p. 185-204, 2010.

ROOY, M. P. VAN. *A rainfall anomaly index independent of time and space*. Notos, Pretoria, v. 14, p. 43-48, 1965.

ROSSI, G.; BENEDINI, M.; TSAKIRIS, G.; GIAKOUMAKIS, S. *On regional drought estimation and analysis*. Water Resources Managements, 6, p.249 – 277, 1992.

SALAS, J. D.; Fu, C.; CANCELLIERE, A.; DUSTIN, D.; BODE, D.; PINEDA, A.; VINCENT, E. *Characterizing the severity and risk of drought in the Poudre River*. Colorado. J. Water Resources Planning Management - ASCE 131 (5), p.383 – 393, 2005.

SALVADOR, M. A.; SANTOS, L. S. F. C. *Análise da Precipitação na Cidade de São Paulo no Período 1961-2009*. Revista Brasileira de Climatologia, v. 7, p. 8-19. 2010.

SCHAFER, J.L., GRAHAM, J.W. *Missing data: Our view of the state of the art*. Psychological Methods, 7, 147-177. 2002.

SHAFER, B. A.; DEZMAN, L. E.; *Development of a surface water supply index (SWSI) to Assess the Severity of drought conditions in snowpack runoff areas*. Preprints, Western Snow Conference, Reno, NV, Colorado State University, pp. 164–175. 1982.

SILVA, V. P. R. *On climate variability in Northeast of Brazil*. Journal of Arid Environments. v. 58, p. 575-596, 2004.

SILVA, B. F. P. *Contribuição dos distúrbios ondulatórios de leste para a chuva no leste do Nordeste do Brasil: evolução sinótica média e simulações numéricas*. Dissertação de mestrado em Meteorologia – Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo. Disponível em <http://www.teses.usp.br/teses/disponiveis/14/14133/tde-04102011-221401/>. Acesso em: 07/03-2013, 2011.

SILVA, V. M. A.; PATRÍCIO, M. C. M.; RBEIRO, V. H. A.; MEDEIROS, R. M. *O desastre da seca no Nordeste brasileiro*. Polêmica, v.12, n.2, p.284-293, 2013.

STRANG, D. M. *Utilização dos Dados Pluviométricos de Fortaleza, CE, Visando Determinar Probabilidades de Anos Secos e Chuvosos*. São José dos Campos, CTA, Relatório Técnico ECA - 03/79, 1979.

SUASSUNA, João. *Uma metodologia para a preparação do combate aos efeitos da seca*. Disponível em: <<http://www.fundaj.gov.br/docs/tropico/desat/js0507799.htm>>. Acesso em: 14 set. 2011.

SUPERINTENDÊNCIA DO DESENVOLVIMENTO DO NORDESTE – SUDENE. Disponível em <http://www.sudene.gov.br/site/extra.php?cod=108idioma=ptbr>. Visitado em 08/11/2012.

THOM, H.C.S. *A note on the gamma distribution*. Monthly Weather Review, Washington, v86, 1958, p.117-122.

UNCCD (United Nations Convention to Combat Drought and Desertification). *United Nations Convention to Combat Drought and Desertification in Countries Experiencing Serious Droughts and or Desertification*. Particularly in Africa. Paris, 1994.

- USDA. *Major world crop areas and climatic profiles*. World Agricultural Outlook Board, US department of Agriculture, Agricultural Handbook. n. 664, p.157 – 170, 1994.
- VAREJAO-SILVA, M. *A Meteorologia e climatologia*. Recife, 2013.
- VILLA, M. A. *Vida e morte no sertão: história das secas no Nordeste nos séculos XIX e XX*. Revista Brasileira de História. São Paulo, v.22, n.43, p.251-254. 2002.
- WALPOLE, R. E.; MYERS, R. H. *Probability and statistics for engineers and scientists*. New York: Macmillan, 1978. 580 p.
- WILHITE, D. A.; GLANTZ, M. H. *Understanding the drought phenomenon: the role of definitions*. Water International, 10, p.111 – 120. 1985.
- WILHITE, D.A. Drought. In: HOLTON, J.A; PYLE, C.; CURRY, J.A. *Encyclopedia of Atmospheric Science*. New Yourk: Elsevier Science, 2003.
- WILHITE, D. A.; SVOBODA, M. D.; HAYES, M. J. *Understanding the complex impacts of drought: A key to improving drought mitigation and preparedness*. Water Resources Management, 21 (5), p. 763-774. 2007.
- YOO, C.; KIM, S. *EOF analysis of surface soil moisture field variability*. Advanced Water Resources, n.27, p.831 – 842, 2004.
- YOU, J. K., HUBBARD, G., GODDARD, S. *Comparison of methods for spatially estimating station temperatures in a quality control system*. International Journal of Climatology, 28, 777-787, DOI: joc.1571. 2007.
- YUE, S. *The Gumbel mixed model applied to storm frequency analysis*. Water Resources Managements, 14, p.377 – 389, 2000.
- YUE, S., RASMUSSEN, P. *Bivariate frequency analysis: discussion of some useful concepts in hydrological application*. Hydrological Process, n.16, p.2881–2898, 2002.
- YUE, S.; OUARDA, T. B. M. J.; BOBÉE, B.; LEGENDRE, P.; BRUNEAU, P. *The Gumbel mixed model for flood frequency analysis*. Journal of Hydrology, n.226, p.88 – 100, 1999.
- ZHANG, L.; SINGH, V. P. *Bivariate flood frequency analysis using the copula method*. Journal Hydrologic Engineering (ASCE). n.11 (2), p.50 – 164, 2006.
- ZENG, N. *Drought in the Sahel*. Science, n. 302, p. 999–1000, 2003.
- ZHANG, Q. *Drought and its impacts*. In: Chen, H. (Ed.), China Climate Impact Assessment. China Meteorological Press, Beijing. 2003.
- ZOU, X.; PANMAO, Z.; QIANG, Z. *Variations in droughts over China: 1951–2003*. Geophysical Research Letters, v.32, I. 4. p.45 – 67.

APÊNDICE

APÊNDICE A - Estatísticas descritiva do Estado de Pernambuco

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	23,5	132,5	66,6	831,6	28,8	0,4	0,6	-0,4
Fev	39,8	121,8	70,0	507,1	22,5	0,3	0,5	-0,9
Mar	48,1	186,6	102,5	946,2	30,8	0,3	0,4	0,0
Abr	52,5	194,0	93,1	601,8	24,5	0,3	1,5	4,3
Mai	15,9	155,1	70,0	1093,8	33,1	0,5	0,7	0,3
Jun	7,3	211,5	63,7	2663,0	51,6	0,8	1,0	0,3
Jul	2,8	207,6	55,2	2398,0	49,0	0,9	1,1	0,7
Ago	1,2	121,8	31,2	991,2	31,5	1,0	1,2	0,8
Set	1,4	72,5	14,8	255,1	16,0	1,1	1,9	3,9
Out	3,5	41,7	14,9	61,2	7,8	0,5	1,5	2,4
Nov	6,0	46,3	19,2	86,5	9,3	0,5	0,6	-0,1
Dez	9,1	79,4	36,6	294,7	17,2	0,5	0,6	-0,7

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE B - Estatísticas descritiva do Estado do Ceará

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	68,3	189,7	110,2	924,0	30,4	0,3	0,8	0,1
Fev	74,2	205,5	124,6	1050,3	32,4	0,3	0,8	0,1
Mar	130,2	310,1	201,2	1731,4	41,6	0,2	0,5	0,0
Abr	114,1	294,0	187,9	1583,3	39,8	0,2	0,8	0,6
Mai	36,9	175,8	91,6	766,1	27,7	0,3	0,8	1,0
Jun	4,6	130,5	37,3	473,8	21,8	0,6	2,2	7,6
Jul	5,8	68,4	16,8	127,6	11,3	0,7	2,8	10,4
Ago	0,9	18,0	4,8	12,0	3,5	0,7	1,9	4,1
Set	0,0	11,7	2,7	7,5	2,7	1,0	1,7	2,9
Out	0,0	24,7	6,2	49,4	7,0	1,1	1,3	0,5
Nov	0,4	40,0	8,0	80,1	8,9	1,1	2,0	3,6
Dez	7,1	94,7	31,5	459,5	21,4	0,7	1,1	0,5

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE C - Estatísticas descritiva do Estado da Paraíba

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	37,1	158,7	90,6	1442,5	38,0	0,4	0,2	-1,3
Fev	46,5	159,9	99,0	1326,0	36,4	0,4	0,1	-1,4
Mar	64,7	233,1	150,7	2787,9	52,8	0,4	-0,2	-1,3
Abr	46,3	240,2	139,2	2629,1	51,3	0,4	-0,1	-1,1
Mai	40,8	154,3	86,4	700,0	26,5	0,3	0,5	0,2
Jun	18,2	203,2	56,2	1478,9	38,5	0,7	2,2	4,6
Jul	10,2	176,2	40,7	1517,6	39,0	1,0	2,1	3,4
Ago	2,5	117,8	22,2	672,8	25,9	1,2	2,1	4,0
Set	0,5	49,3	8,3	116,2	10,8	1,3	2,3	4,7
Out	2,8	22,4	10,9	22,6	4,8	0,4	0,4	-0,5
Nov	1,9	28,7	12,1	45,3	6,7	0,6	0,8	0,0
Dez	10,9	63,7	31,9	234,1	15,3	0,5	0,6	-0,8

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE D - Estatísticas descritiva do Estado de Minas Gerais

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	103,1	292,2	139,1	1813,1	42,6	0,3	2,4	5,8
Fev	51,4	148,2	81,8	479,6	21,9	0,3	1,2	1,7
Mar	81,7	149,2	115,9	268,2	16,4	0,1	0,0	-0,2
Abr	29,8	71,0	44,3	140,9	11,9	0,3	0,8	-0,5
Mai	2,8	33,6	17,2	90,1	9,5	0,6	0,3	-1,2
Jun	0,7	22,4	8,7	59,3	7,7	0,9	0,7	-1,1
Jul	0,2	29,6	8,8	76,3	8,7	1,0	1,1	0,1
Ago	1,1	22,9	8,8	45,3	6,7	0,8	0,9	-0,6
Set	9,5	50,6	21,4	124,0	11,1	0,5	1,2	0,6
Out	40,0	99,2	66,2	259,8	16,1	0,2	0,4	-0,7
Nov	126,1	215,5	168,1	497,0	22,3	0,1	0,3	-0,4
Dez	157,4	291,3	199,5	1208,9	34,8	0,2	1,0	0,4

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE E - Estatísticas descritiva do Estado da Bahia

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	20,7	211,9	89,4	872,4	29,5	0,3	0,6	2,0
Fev	25,4	116,4	76,7	431,6	20,8	0,3	-0,3	-0,5
Mar	28,3	173,7	83,5	1002,1	31,7	0,4	0,5	-0,2
Abr	17,9	131,2	53,0	473,8	21,8	0,4	1,3	1,9
Mai	3,1	159,6	38,0	1089,7	33,0	0,9	1,6	3,2
Jun	0,0	167,8	36,4	1537,9	39,2	1,1	1,4	1,9
Jul	0,0	151,7	30,4	1126,7	33,6	1,1	1,5	2,6

Ago	0,0	125,5	25,4	729,1	27,0	1,1	1,4	2,3
Set	1,0	86,9	19,3	275,5	16,6	0,9	1,8	4,2
Out	3,9	79,5	39,6	268,9	16,4	0,4	0,3	-0,4
Nov	2,6	183,1	85,1	2025,0	45,0	0,5	0,2	-0,7
Dez	1,6	174,6	80,6	1675,8	40,9	0,5	0,1	-0,7

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE F - Estatísticas descritiva do Estado do Piauí

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	31,9	108,6	61,7	261,1	16,2	0,3	0,7	0,9
Fev	63,5	136,0	99,1	419,4	20,5	0,2	-0,1	-1,2
Mar	100,1	247,1	176,3	1033,4	32,1	0,2	-0,2	0,1
Abr	113,3	242,9	174,8	1174,0	34,3	0,2	0,0	-0,9
Mai	36,3	148,6	93,6	756,5	27,5	0,3	-0,3	-0,6
Jun	17,0	175,8	46,6	980,8	31,3	0,7	2,9	9,0
Jul	16,0	158,8	40,1	915,9	30,3	0,8	2,6	6,9
Ago	4,9	100,6	17,7	365,7	19,1	1,1	3,3	11,0
Set	0,8	29,4	7,1	53,8	7,3	1,0	1,6	1,6
Out	0,0	12,3	2,9	8,5	2,9	1,0	1,7	2,7
Nov	0,7	23,2	6,5	22,3	4,7	0,7	1,6	3,6
Dez	10,6	51,6	22,9	80,6	9,0	0,4	1,6	2,6

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE G - Estatísticas descritiva do Estado do Rio Grande do Norte

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	31,9	108,6	61,7	261,1	16,2	0,3	0,7	0,9
Fev	63,5	136,0	99,1	419,4	20,5	0,2	-0,1	-1,2
Mar	100,1	247,1	176,3	1033,4	32,1	0,2	-0,2	0,1
Abr	113,3	242,9	174,8	1174,0	34,3	0,2	0,0	-0,9
Mai	36,3	148,6	93,6	756,5	27,5	0,3	-0,3	-0,6
Jun	17,0	175,8	46,6	980,8	31,3	0,7	2,9	9,0
Jul	16,0	158,8	40,1	915,9	30,3	0,8	2,6	6,9
Ago	4,9	100,6	17,7	365,7	19,1	1,1	3,3	11,0
Set	0,8	29,4	7,1	53,8	7,3	1,0	1,6	1,6
Out	0,0	12,3	2,9	8,5	2,9	1,0	1,7	2,7
Nov	0,7	23,2	6,5	22,3	4,7	0,7	1,6	3,6
Dez	10,6	51,6	22,9	80,6	9,0	0,4	1,6	2,6

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE H - Estatísticas descritiva do Estado de Alagoas

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	17,1	45,0	32,0	69,3	8,3	0,3	-0,3	-1,0
Fev	20,0	84,7	39,5	222,6	14,9	0,4	1,4	2,2
Mar	25,1	123,9	62,0	462,5	21,5	0,3	1,1	1,9
Abr	21,2	110,8	77,3	568,2	23,8	0,3	-0,6	-0,3
Mai	15,2	168,4	102,5	1395,4	37,4	0,4	-0,5	-0,3
Jun	14,7	182,7	105,5	1874,8	43,3	0,4	-0,1	-0,6
Jul	17,0	210,5	109,6	2440,4	49,4	0,5	0,2	-0,6
Ago	10,7	110,7	61,9	687,4	26,2	0,4	-0,1	-0,7
Set	10,1	61,6	34,4	268,9	16,4	0,5	0,2	-1,2
Out	4,5	39,3	22,9	113,3	10,6	0,5	-0,1	-1,3
Nov	4,4	40,6	20,5	58,9	7,7	0,4	0,6	1,2
Dez	8,0	53,7	29,6	130,5	11,4	0,4	-0,1	-0,4

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

APÊNDICE I - Estatísticas descritiva do Estado de Sergipe

Meses	Mínimo	Máximo	Média	Variância	DP	CV	CA	CC
Jan	24,7	47,5	36,0	40,5	6,4	0,2	0,2	-0,7
Fev	30,4	54,1	43,0	34,9	5,9	0,1	-0,2	-0,3
Mar	58,5	96,8	73,8	124,2	11,1	0,2	0,5	-0,9
Abr	66,7	139,9	105,5	465,0	21,6	0,2	0,0	-0,8
Mai	68,3	199,5	128,7	1189,7	34,5	0,3	0,0	-0,7
Jun	57,6	188,1	117,0	1420,9	37,7	0,3	0,2	-0,8
Jul	60,0	169,7	112,5	1162,3	34,1	0,3	0,0	-1,1
Ago	25,3	101,8	65,7	512,4	22,6	0,3	-0,3	-0,9
Set	19,2	70,0	47,3	285,2	16,9	0,4	-0,3	-1,2
Out	9,4	41,1	27,7	79,7	8,9	0,3	-0,5	-0,4
Nov	11,2	44,3	28,1	79,8	8,9	0,3	0,1	-0,9
Dez	18,9	50,4	39,0	95,8	9,8	0,3	-0,5	-1,0

DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose. Fonte: Elaborada pelo autor.

