

ESTUDO DA PRECIPITAÇÃO DO SEMIÁRIDO BAIANO USANDO ANÁLISES DE AGRUPAMENTO E DISTRIBUIÇÃO DE PROBABILIDADE

STUDY OF BAITIAN SEMIARIDUM PRECIPITATION USING GROUP ANALYSIS AND DISTRIBUTION OF PROBABILITY



Resumo:

A precipitação é de soberana importância aos estudos climáticos, pois quando determinada em excesso ou deficiência, provocam calamidades aos setores: agropecuário, agronegócio, irrigação, geração de energia, armazenamento, represamento e abastecimento humano e socioeconômico entre outros. Utilizaram-se os dados pluviométricos mensais de 91 postos no período de 1984 a 2014 (31 anos). Os dados tratados foram submetidos a análise de agrupamento e os mesmos geraram três regiões homogêneas de estudo. O objetivo deste trabalho foi testar o ajuste de diferentes funções de distribuição de frequência à precipitação média anual. Foram testadas as distribuições de probabilidade Exponencial, Gumbel I, Gama, Log-normal, Normal, Weibull, Logística, Beta entre outras. A série de dados utilizados foram obtidos da base de dados da Superintendência do Desenvolvimento do Nordeste (SUDENE) e Agência Nacional das Águas (ANA). A aderência das distribuições foi testada pelo teste de Kolmogorov-Smirnov ao nível de 5%. A distribuição que obteve melhor desempenho nos ajustes à precipitação de do Semiárido baiano para a série mensal foi a Logística, com p-valor a 0,8848 e com o nível de significância $\alpha=0,05$. Os ajustes foram avaliados em gráficos. As distribuições Gama 2 e GEV também apresentaram bons resultados em relação ao estudo.

Abstract:

Rainfall is of sovereign importance for climatic studies, as it is determined in excess or deficiency, causing calamities to the sectors: agriculture, agribusiness, irrigation, power generation, storage, damming and human and socioeconomic supply among others. The monthly rainfall data of 91 stations during the period 1984 to 2014 (31 years) were used. The treated data were submitted to cluster analysis and they generated three homogeneous regions of study. The objective of this work is to test the adjustment of different frequency distribution functions to the press. They were tested as distributions of probability Exponential, Gumbel I, Range, Log-normal, Normal, Weibull, Logística, Beta among others. A series of data used in the database of the Northeast Development Superintendency (SUDENE) and the National Water Agency (ANA) An. The adjustments were evaluated in graphs. The Gama 2 and GEV distributions also showed good results in relation to the study adherence of the distributions was tested by the Kolmogorov-Smirnov test at the 5% level. The distribution that obtained the best performance in the adjustments to the propagation of the Semi-arid Baiano for monthly series for Logistics, with p-value at 0.8848 and with the level of significance $\alpha = 0.05$

Valneli da Silva Melo¹
Kely Dayane Silva do Ó¹
Milena Pereira Dantas²
Raimundo Mainar de Medeiros³

¹Doutorandas em Engenharia Ambiental e Sanitária, UEPB, Campina Grande, PB, valnelismello@hotmail.com, kely.dayane@hotmail.com; ²Mestra em Meteorologia, UFCG, Campina Grande, PB, milenapdantass@gmail.com; ³Dr. em Meteorologia e Pesquisador da Universidade Federal Rural de Pernambuco, UFRPE, PE, Brasil, e-mail: mainarmedeiros@gmail.com

Contato principal

Valneli da Silva Melo¹



Palavras chave: análise estatísticas, níveis de probabilidade, regimes pluviométricos,

Keywords : Statistical analysis, probability levels, rainfall regimes.



INTRODUÇÃO

A água é um recurso essencial para a manutenção da “água doce”, este fator encontra-se atrelado às múltiplas atividades desenvolvidas por meio deste recurso, entre elas, abastecimento para consumo humano, atividades industrial e agrícola, e sendo de importância para os ecossistemas conforme Rebouças (2006).

O monitoramento da precipitação pluvial é ferramenta indispensável na mitigação de secas, cheias, enchentes, inundações, alagamentos conforme ressaltam Paula et al. (2010). Dentre os elementos do clima de áreas tropicais, a precipitação é o que mais influencia a produtividade agrícola em conformidade com os autores Ortolani e Camargo (1987), principalmente nas regiões semiárida, onde o regime de chuvas é caracterizado por eventos de curta duração e alta intensidade (Santana et al. 2007). Em função disso a sazonalidade da precipitação concentra quase todo o seu volume durante os cinco a seis meses no período chuvoso em conformidade com Silva (2004).

Segundo Santos et al (2010) a caracterização das chuvas intensas é imprescindível para solucionar problemas de interesse da engenharia, de modo especial o controle do escoamento superficial em áreas urbanas e rurais. O processo erosivo e sua intensidade dependem principalmente das condições climáticas da região, fatores relacionados à topografia, cobertura do solo e às propriedades do mesmo (GONÇALVES,2002).

Melo et al. (2015) mostraram que são importantes os planejamentos urbano, rural e ambiental, bem como o armazenamento de água para múltiplos fins. O trabalho identificou variações e tendências mensais e anuais das variáveis meteorológicas: temperaturas máxima e mínima do ar, precipitação pluvial, umidade relativa do ar, número de dias com chuva e insolação total no município de Bom Jesus – PI. Os autores utilizaram as seguintes funções de distribuição teórica de probabilidade: Weibull, Log-normal e Logística para o ajuste dos valores das variáveis supracitadas. Utilizaram o teste Kolmogorov-Smirnov (KS) para verificar o ajuste das funções teóricas. Os resultados mostraram aumentos da temperatura máxima e redução da temperatura mínima, condição importante para o processo de desertificação na área estudada. Essa condição provoca estresse hídrico nas culturas agrícolas e em consequência baixa produção.

No Nordeste Brasileiro (NEB), verifica-se ao longo do ano um período curto de 3 a 4 meses com precipitações e um período longo sem a sua ocorrência, geralmente chamado de período de estiagem, exibindo alta capacidade de evapotranspiração durante o ano, caracterizando um clima semiárido. O semiárido nordestino se destaca pelas precipitações médias anuais muito irregulares e com grande variabilidade espacial. As precipitações médias variam entre 200 a 700 mm ano⁻¹, quando comparadas com outras regiões semiáridas do mundo, este índice pluviométrico não é tão baixo, no entanto as temperaturas são elevadas, as perdas por evapotranspiração são

acentuadas e o semiárido com maior densidade populacional do mundo evidenciando a necessidade, do ponto de vista social, de estudar formas melhores de convivência com a escassez hídrica (CABRAL et al, 2007).

Vários estudos têm sido evidenciados tendo como tema a variabilidade e suas características climáticas, portanto ainda está longe de chegar a um consenso, principalmente, no que tange a influência local nas tendências globais. Contudo os estudos de variabilidade climática local, principalmente sobre a temperatura, umidade relativa do ar e precipitação são essenciais para se compreender os impactos que estas variações podem trazer para a população em geral. Se realmente as flutuações forem significativas em âmbito local, essas mudanças poderão causar impactos socioeconômicos, ambientais e sociais (proliferação de vetores de doenças), (MALUF et. al., 2011).

Segundo Bergamaschi et al (2009) o clima é o principal fator responsável pelas oscilações e frustrações das safras agrícolas brasileiras. Quando se trata da região NEB, principalmente nas épocas das secas prolongadas, são as maiores responsáveis pela variabilidade da produção agrícola.

Na hidrologia, o uso de técnicas de mineração de dados no estudo do comportamento espaço temporal de variáveis meteorológicas, em especial, a precipitação, pode ser encontrado em diversos estudos. Souza et al. (2012) utilizaram técnicas de agrupamento hierárquico para análise da precipitação e temperatura no estado do Mato Grosso do Sul. Boschi et al. (2011) utilizaram o algoritmo k-means na determinação do comportamento espacial e temporal da precipitação no estado do Rio Grande do Sul, aplicando a metodologia a dois decênios subsequentes, encontrando alteração no volume de precipitação entre os decênios. Conde et al. (2012), Silva et al. (2010) e Dourado et al. (2010) definiram regiões de precipitação homogêneas através do algoritmo k-means.

Objetiva-se determinar regiões homogêneas, descrever o comportamento da série temporal da precipitação, verificar quais das principais distribuições de probabilidade se ajustam aos dados da série de precipitações pluviométricas, usar os testes de aderência para verificar quais distribuições melhor ajustou aos dados em análise e determinar qual função de probabilidade das funções testadas melhor representa a precipitação pluvial do Semiárido da Bahia.

MATERIAIS E MÉTODOS

CARACTERIZAÇÃO DA ÁREA DE ESTUDO

O Estado da Bahia limita-se com Alagoas, Sergipe, Pernambuco, além do Piauí ao norte, Minas Gerais e Espírito Santo ao sul, Goiás e Tocantins a oeste e ao leste com o oceano Atlântico. O estado ocupa área de 564.830,859 km², no interior do território existem

diferentes tipos de relevos, climas, vegetações e hidrografia.

No território baiano o clima que predomina é o tropical, devido a isso as temperaturas são elevadas praticamente o ano todo, superior aos 26°C, exceto na serra do Espinhaço, onde as temperaturas são amenas. O clima semiárido se apresenta no sertão do Estado com pouca precipitação, ocasionando longos períodos de estiagem. Na região de Ilhéus os índices pluviométricos podem

superar os 1.500 mm e no sertão não ultrapassa 500 mm ano⁻¹.

As chuvas são irregulares, ocasiona a ocorrências de grandes períodos de seca. Na Figura 1 pode ser visto a distribuição espacial dos postos pluviométricos na área de estudo e o relevo do estado. Posteriormente, esses dados foram tratados e organizados para se obter as estatísticas básicas para cada um desses postos.

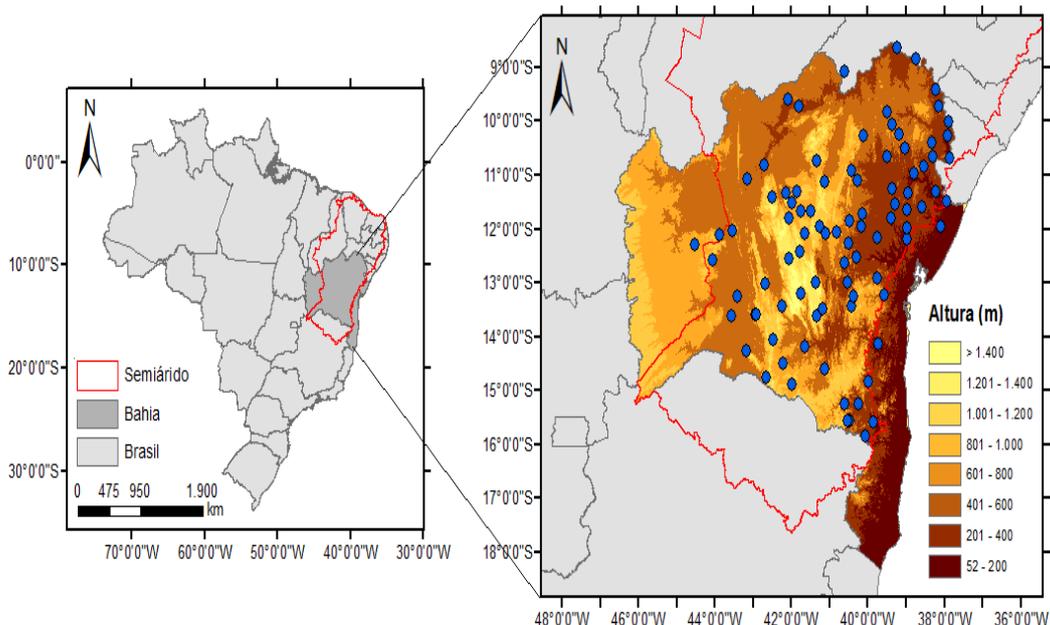


Figura 1. Posição geográfica do estado da Bahia e distribuição espacial dos postos pluviométricos.

O relevo (Figura 1) é marcado pela presença de planícies, planaltos e depressões, e as formaturas tabulares e planas. O ponto mais alto da Bahia é representado pelo Pico das Almas, situado na Serra das Almas. Ao oeste, observam-se áreas de chapadas. Os planaltos ocupam quase todo o estado, porém o relevo que predomina no estado baiano é a depressão. As planícies estão situadas na região litorânea, altitude abaixo de 200 m. O volume de chuvas de uma localidade pode ser intensificado ou reduzido conforme as características do relevo.

DADOS DE PRECIPITAÇÃO

Os dados utilizados nesse estudo foram cedidos pela Superintendência do Desenvolvimento do Nordeste (SUDENE) e Agência Nacional de Águas (ANA), as estações pluviométricas foram selecionadas de modo a se obter o maior período de dados mensais sem falhas e uma distribuição homogênea dentro da área de estudo, totalizando 91 postos pluviométricos, com tamanho amostral superior a 31 anos (1984 a 2014), foram calculadas várias estatísticas descritivas básicas (por exemplo, média, valor máximo, mínimo, variância, coeficiente de assimetria). Para a confecção dos mapas, utilizou-se de software de geoprocessamento QGIS. Para a

análise de cluster (agrupamento) os dados foram consistidos, preenchendo-se eventuais falhas e realizado sua homogeneização. Para definição do número de grupos foi aplicado preliminarmente o método de agrupamento hierárquico depois se ajustou as funções de distribuições de probabilidade para verificar qual apresenta melhor ajuste.

DELIMITAÇÕES DAS SUB-REGIÕES POR ANÁLISE DE AGRUPAMENTO

A análise de agrupamento (AA) é ferramenta útil para análise de dados pluviométricos. 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 estrutura num conjunto de dados homogêneos que têm de ser separados.

AA, utilizada em processos de classificação, consistem em determinar o nível de similaridade ou dissimilaridade entre indivíduos, aplicando função de agrupamento a determinada variável, transformando um conjunto heterogêneo de unidades, neste caso estações

meteorológicas, em grupos caracterizados pela homogeneidade interna e encontrando-se, também, grande número pela heterogeneidade externa (JOHNSON, WICHERN, 2007).

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

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

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 à 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 utilizado quando se têm muitos objetos para agrupar, com pequenas variações. O critério utilizado de homogeneidade 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 o menor SQD dentro os 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, 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 municípios com precipitação similar, aplicou-se novamente o método de Ward para determinar, dentro de cada período mensal definido, 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:

$$H = \frac{V - \mu_v}{\sigma_v} \quad (1)$$

e

$$V = \sqrt{\left\{ \frac{\sum_{i=1}^N n_i (\tau^i - \tau^R)}{\sum_{i=1}^N n_i} \right\}} \quad (2)$$

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.(2) 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 $H > 2,0$.

DISTRIBUIÇÕES DE PROBABILIDADE

Um conjunto de dados meteorológico e/ou hidrológicos 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 hidrológicos é conhecido, no âmbito da hidrologia ou da meteorologia, como série histórica e

consiste, basicamente, de uma amostra extraída de uma população.

Com base nesta amostra, podemos calcular alguns indicadores e medidas estatísticas importantes, como média, desvio padrão (variância), assimetria, curtose e distribuição de frequência dos dados observados na amostra. E com isso, modelar esta distribuição de frequência com base em modelo matemático, constituído de parâmetros, conhecido como Distribuição de Probabilidades.

Após as análises descritivas as variáveis média mensal da série de precipitações pluviométricas foram submetidas ao ajuste das seguintes distribuições: Normal, Exponencial, Gama com 1,2 e 3 parâmetros, Gumbel, LogNormal, Logística, Weibull e outras, utilizando as metodologias descritas em Mayer (2001), Walpole (2009) e Hann (1994). Os ajustes foram feitos nos softwares Estatística 8 e Excel 2010, utilizando a metodologia de máxima verossimilhança.

A verificação da qualidade das distribuições de probabilidades em explicar o comportamento dos dados observados foi realizada por meio dos testes de aderência. Na literatura estatística, conforme se encontra em Campos (1976), os principais testes de aderência utilizados são: qui-quadrado, Kolmogorov-Smirnov e Lilliefors, neste usou-se o teste de aderência de Kolmogorov-Smirnov.

TESTE DE KOLMOGOROV-SMIRNOV

De acordo com Campos (1976) Kolmogorov-Smirnov é um teste que avalia a concordância entre a distribuição observada da amostra e uma determinada distribuição teórica. Avalia se os dados amostrais se aproximam razoavelmente de uma determinada distribuição. Para isso utilizamos a função distribuição acumulada observada, compara-se com a teórica, determina-se o ponto em que estas distribuições mais divergem, e testamos se essa divergência é aleatória ou não.

O teste exige que os dados devam seguir ao menos uma escala ordinal.

Seja $F_0(X)$ uma distribuição teórica acumulada e $S_n(X)$ uma distribuição observada em uma amostra de n observações (distribuição empírica). Encontra-se a seguir o maior valor das diferenças entre $F_0(X)$ e $S_n(X)$, ou seja,

$$D = \max |F_0(X) - S_n(X)|$$

Compara-se o valor observado com o valor crítico (tabelado). Na tabela associa-se o valor observado com o seu valor-p.

RESULTADOS E DISCUSSÃO

Com base nos totais mensais precipitados dos postos pluviométricos e na análise de agrupamento foram obtidas três sub-regiões pluviometricamente homogêneas e distintas entre si. A Figura 2 exibe as três sub-regiões, encontradas na análise de agrupamento, onde as regiões apresentam características pluviométricas similares entre si.

Primeiro grupo (sub-região 1) comportam as estações no sul, seu total médio anual de precipitação varia em torno de 676,5 mm, sendo os meses de novembro a janeiro os mais chuvosos e agosto o mais seco. Segundo grupo (sub-região 2) contempla a parte central da área de estudo, nessa área o total médio anual é o mais alto chegando em média a 1073,5 mm, sendo seu trimestre mais chuvoso nos meses de março a junho e o terceiro grupo comportam as estações da área norte da região com total médio anual de 552,3 mm com o trimestre mais chuvoso nos meses de novembro a janeiro.

O agrupamento tem boa concordância com peculiaridades físico-geográficas e topográficas da região de estudo e com efeito de continentalidade. A análise das normas mensais mostra que dentro dos grupos estudados, o segundo apresenta as maiores normas mensais de chuva. Aparentemente isto associado com efeito orográfico. Braga et al. (1998) afirmam que o Estado da Bahia, devido a sua posição geográfica, seu território é influenciado por sistemas atmosféricos de várias escalas, em diferentes épocas do ano, fazendo com que, haja significativa variação de precipitação ao longo do seu território.

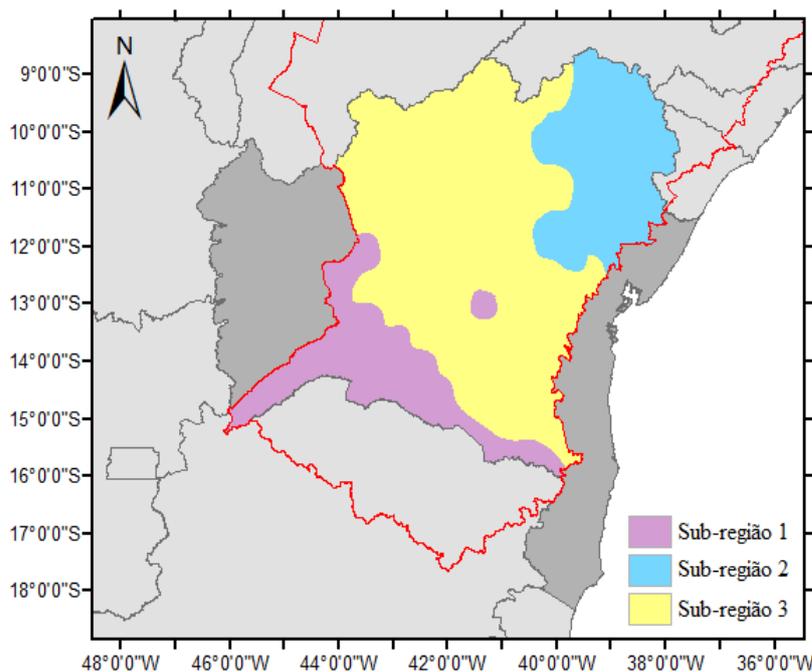


Figura 2. Regiões homogêneas do semiárido baiano

.Com base na análise de agrupamento dos dados de precipitação da área estudada, tem-se a estatística descritiva, observa-se na Tabela 1 a distribuição mensal da precipitação para a região. Observa-se que há predominância do período chuvoso entre os meses de novembro a janeiro, essa variabilidade interanual da pluviometria nesta região está associada a 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).

A análise descritiva dos dados objetivou explorar visualmente as informações da série para detectar comportamentos e verificar a presença de observações atípicas, ou seja, esta etapa teve a finalidade de conhecer as variáveis analisadas. Autores como Bussab e Morettin (2002) e Triola (2009) enfatizam a necessidade e a importância da análise descritiva das variáveis analisadas antes de qualquer análise inferencial, pois, fatos como a presença de valores atípicos, o tipo de comportamento da variável analisada e até mesmo os erros de digitação em bancos de dados, podem distorcer os resultados da análise inferencial, provocando conclusões incorretas ou não precisas.

Todas as análises descritivas foram realizadas nos programas Excel 2010, seguindo procedimentos básicos descritos em Bussab e Morettin (2002).

A Tabela 1, apresenta as análises descritivas dos dados

onde os meses em estudo apresentaram coeficiente de assimetria (CA) positivo, com assimetria à direita ($CA > 0$). Os coeficientes de variação (CV) ($CV > 20\%$) indicam que há alta dispersão entre os valores médios de precipitação, ou seja, há grande variabilidade da chuva na região do Semiárido baiano. Com relação à curtose, todos os valores médios mensais apresentaram coeficientes de curtose positivos e negativos, indicando que a distribuição dos dados é do tipo leptocúrtica e platicúrtica.

Ainda levando em consideração a análise de agrupamento, tem-se na Figura 3, o perfil pluviométrico das sub-regiões homogêneas, onde observa-se que a sub-região 2 apresenta um maior índice de precipitação com variações máximas nos meses de maio a julho, já a sub-região 1 apresenta os menores valores de precipitação, tendo nos meses de maio a setembro valores mínimos atingidos por essa sub-região, com precipitações máximas nos meses de novembro a janeiro, o que já é esperado nessa região. A sub-região 3, apresenta um perfil pluviométrico sem extremos, onde pode-se observar essa região apresenta pouca variação durante todos os meses.

Observa-se na Figura 4 que 35 postos as precipitações anuais fluíram abaixo da normal climatológica, sete postos oscilaram próximo a média e quarenta e três postos apresentaram médias anuais acima da climatológica. Estas flutuações estão em conformidade com as flutuações dos fenômenos de larga escala El Niño(a), média e mini oscilações locais e regionais que contribuem para as flutuações locais.

ESTUDO DA PRECIPITAÇÃO DO SEMIÁRIDO BAIANO USANDO ANÁLISES DE AGRUPAMENTO E DISTRIBUIÇÃO DE PROBABILIDADE

Tabela 1. Estatísticas descritiva da precipitação do Semiárido baiano.

Meses	Mínimo	Máximo	1°Quartil	Mediana	3°Quartil	Média	DP	CV (%)	CA	CC
Janeiro	20,7	211,9	67,1	90,9	106,7	89,4	29,5	0,3	0,6	2,0
Fevereiro	25,4	116,4	61	76,3	93,2	76,7	20,8	0,3	-0,3	-0,5
Março	28,3	173,7	56,3	82,1	102,4	83,5	31,7	0,4	0,5	-0,2
Abril	17,9	131,2	36,9	47,2	64,9	53,0	21,8	0,4	1,3	1,9
Mai	3,1	159,6	12,5	28,2	53,2	38,0	33,0	0,9	1,6	3,2
Junho	0,0	167,8	2,9	24,5	55,0	36,4	39,2	1,1	1,4	1,9
Julho	0,0	151,7	1,1	23,3	50,7	30,4	33,6	1,1	1,5	2,6
Agosto	0,0	125,5	1,4	22,8	37,5	25,4	27,0	1,1	1,4	2,3
Setembro	1,0	86,9	7,3	14,8	25,7	19,3	16,6	0,9	1,8	4,2
Outubro	3,9	79,5	28,3	37,4	49,7	39,6	16,4	0,4	0,3	-0,4
Novembro	2,6	183,1	50,3	83,1	121,1	85,1	45,0	0,5	0,2	-0,7
Dezembro	1,6	174,6	48,4	86,5	106,4	80,6	40,9	0,5	0,1	-0,7

Legenda: DP= Desvio padrão; CV= Coeficiente de variação; CA= Coeficiente de assimetria; CC= Coeficiente de curtose, todas as unidades são medidas em mm.

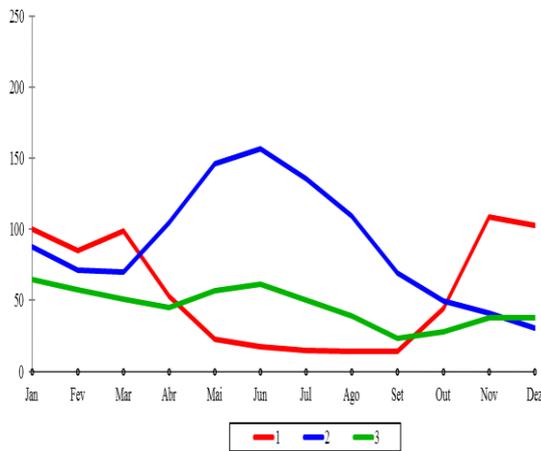


Figura 3. Perfil das sub-regiões homogêneas do Semiárido baiano.

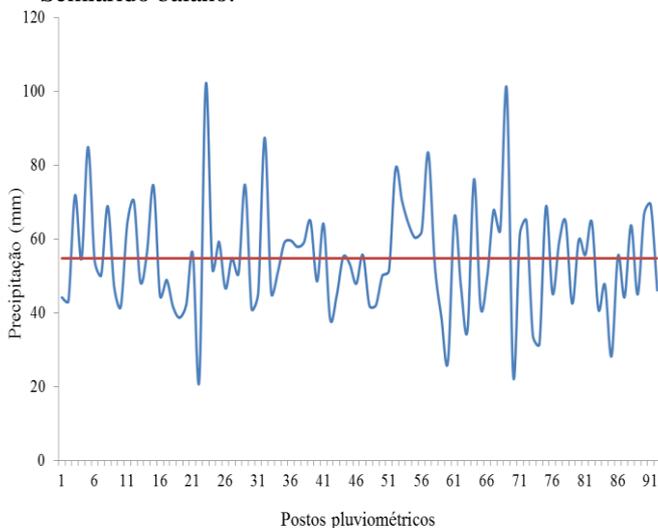


Figura 4. Variabilidade da média anual e a representação histórica dos postos em estudo.

O fato das variáveis apresentarem alta variabilidade indica que a média aritmética pode não representar adequadamente a variável Triola (2009) e também pode ser um primeiro indicativo de que a distribuição de probabilidade normal não é adequada para realizar as estimativas, portanto, conforme argumenta Campos (1976) métodos paramétricos podem não ser adequados na análise inferencial. Após analisar o comportamento pluviométrico da região em estudo, os dados foram ajustados a funções distribuição de probabilidades. A Tabela 2 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 precipitação da região.

Tabela 2. P-valores dos ajustes da precipitação dos municípios do Semiárido baiano as distribuições de probabilidades

Distribuição	p-valor
Beta4	0,1332
Qui-quadrado	0,1577
Erlang	< 0,0001
Exponencial	< 0,0001
Gama (1)	0,0006
Gama (2)	0,6936
GEV	0,7097
Gumbel	< 0,0001
Log-normal	0,5675
Logística	0,8448
Normal	0,2106
Student	< 0,0001
Weibull (1)	< 0,0001

Weibull (2)	0,3419
Weibull (3)	0,3402

A Tabela 2 mostra a proposta do ajuste das distribuições de probabilidades ajustadas aos dados de precipitação da área estudada, observa-se que a distribuição de melhor ajuste foi a distribuição de probabilidade Logística. A distribuição Logística apresentou a melhor bondade de ajuste, com o p-valor calculado igual a 0,8448 que é maior que o nível de significância $\alpha=0,05$, o teste de Kolmogorov-Smirnov é bastante utilizado para análise de aderências de distribuições em estudos climáticos; contudo, o seu nível de aprovação de uma distribuição sob teste é elevado

Estudo realizado por Melo et al (2016) em Parnaçuá – PI, mostram resultados semelhantes onde a distribuição de probabilidade Logística acumulada se adaptou-se ao estudo.

Na Figura 5, observa-se o ajuste da função de probabilidade Logística ajustada aos dados de precipitação, onde pode-se ver que a função obteve um bom ajuste, modelando cerca de 84% dos dados.

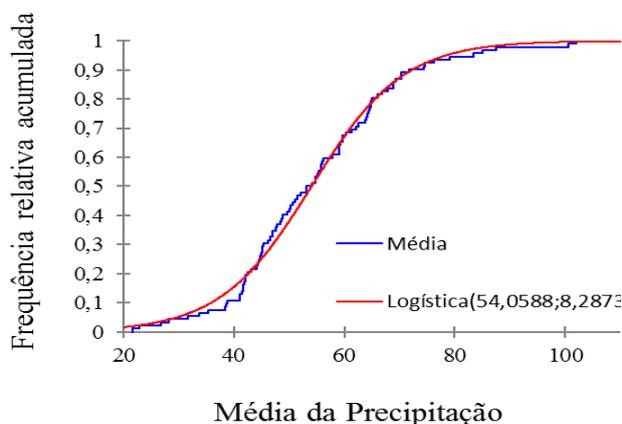


Figura 5. Distribuição de probabilidade Logística acumulada

Observa-se ainda, na Tabela 2, que as distribuições Gama (2) e GEV também apresentam bons ajustes, e que os mesmos também podem representar bem a região de estudo.

Os gráficos representados nas Figuras 6 e 7 respectivamente, representam a qualidade do ajuste da distribuição de probabilidade Gama (2) e GEV aos dados pluviométricos dessa região, onde pode-se ver que os dados foram bem representados por essas distribuições. O teste Kolmogorov-Smirnov usado

nessas análises mostrou que houve bons ajustes de distribuições de probabilidades. Esse teste trabalha diretamente com os dados observados, usando frequências acumuladas.

Moreira et al. (2010) também obteve ajuste usando a distribuição gama para definir diferentes níveis de probabilidade de precipitação na região de Nova Maringá-MT. Murta et al. (2005) também obteve ajustes usando o modelo de distribuição gama para a estimativa da precipitação pluvial mensal em Itapetinga e Vitória da Conquista-BA.

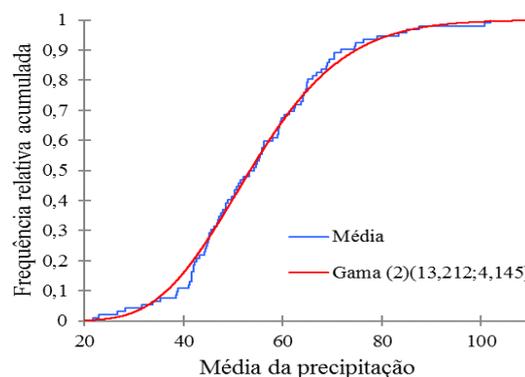


Figura 6. Distribuição de probabilidade Gama acumulada.

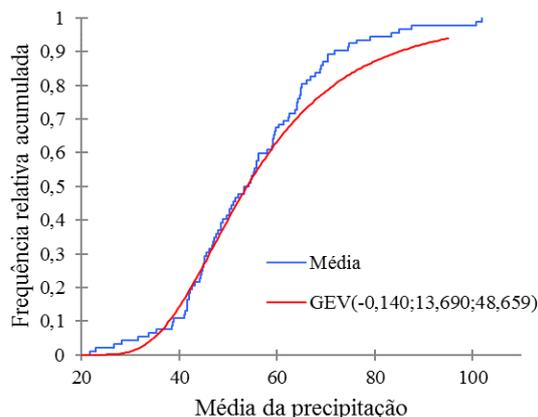


Figura 7. Distribuição de probabilidade GEV acumulada

CONCLUSÕES

A partir das análises feitas na série de precipitações Pluviométricas do semiárido baiano chegou-se as seguintes Conclusões:

As técnicas de agrupamento hierárquico apresentaram bons resultados ao serem aplicadas na identificação de regiões de precipitação homogêneas na área de estudo, onde foram identificadas 3 regiões de precipitações homogêneas, distribuídas de modo coerente com a topografia da região.

Os dados da série de precipitações não seguem a Distribuição normal, não apresenta tendência apenas sazonalidade com periodicidade anual e com alta variabilidade.

O teste de Kolmogorov-Smirnov indicou boa aderência das distribuições Logística, Gama (2) e GEV no nível de 0,05 de Significância. No entanto, a distribuição Logística foi a que Obteve maior aderência. O teste de Kolmogorov-Smirnov indica a rigorosidade do teste e com isso pode-se ver que as distribuições utilizadas nesse teste foram bem representadas.

AGRADECIMENTOS

Os autores deste trabalho agradecem a CAPES pela concessão de bolsa de estudo aos autores.

REFERÊNCIAS BIBLIOGRÁFICAS

- BERGAMASCHI, H; MATZENAUER, R. Milho. In: **Agrometeorologia dos Cultivos: o fator meteorológico na produção agrícola**. MONTEIRO, José Eduardo B. A. (Org.). Brasília, DF: INMET, 2009.
- BOSCHI, R. S.; OLIVEIRA, S. R. D. M.; ASSAD, E. **Técnicas de Mineração de dados para análise da precipitação pluvial decenal no Rio Grande do Sul**. Engenharia Agrícola, v. 31, n. 6, p. 1189– 1201, 2011.
- BRAGA, C.C., de MELO, M.L.D., MELO, E.C.S. **Análise de agrupamento aplicada a distribuição da precipitação no Estado da Bahia**. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 10, Brasília-DF. Anais... Sociedade Brasileira de Meteorologia. p.1857-62, 1998.
- BUSSAB, W. O.; MIAZAKI, E. S.; ANDRADE, D. **Introdução à análise de agrupamentos**. São Paulo: Associação Brasileira de Estatística, 1990. 105p.
- BUSSAB, W. O.; MORETTIN, P. O. **Estatística básica**. 4. ed. São Paulo: Atual, 1987. 320 p.
- CABRAL, J. J. S. P.; SANTOS, S. M. **Água Subterrânea no Nordeste Brasileiro**. In: O uso sustentável dos recursos hídricos em regiões semiáridas. Editora Universitária, Recife – PE, 1 ed., p.65-104. 2007.
- CAMPOS, H. **Estatística experimental não-paramétrica**. 2. ed. Piracicaba: Esalq, 1976.
- CONDE, F. C. et al. **Identificação de Regiões Homogêneas de Precipitação Acumulada Mensal do Distrito Federal via Análise de Cluster**. Anais do XVI Congresso Brasileiro de Meteorologia, 2010.
- DOURADO, C. DA S.; ÁVILA, A. M. H. DE; OLIVEIRA, S. R. M. **Regionalização da precipitação no estado da Bahia por meio de técnicas de mineração de dados**. Anais do XVII Congresso Brasileiro de Meteorologia, 2012.
- EVERITT, B. **Cluster analysis**. New York: Halsted Press, 1991.
- GONÇALVES, J. L. de M.; STAPE, J. L.; WICHERT, M. C. P.; GAVA, J. **Manejo de resíduos vegetais e preparo do solo. Conservação e cultivo de solos para plantações florestais**. Instituto de Pesquisas e Estudos Florestais (IPEF). Cap. 3, p. 133 - 204, Piracicaba, São Paulo, 2002.
- HAAN, C. T. **Statistical Methods in Hydrology**. 6. ed. Ames: Iowa University Press, 1994. 378p.
- HASTENRATH, S., 1984: **Interannual variability and annual cycle: mechanisms of circulation and climate in the tropical Atlantic**. Mon. Wea. Rev., **112**, 1097-1107.
- JOHNSON, R. A., e WICHERN, D. W. **Applied multivariate statistical analysis**. Prentice Hall, 792 pp, 2007.
- MALUF, R. S.; ROSA, T. S. **Mudanças climáticas desigualdade sociais e populações vulneráveis no Brasil**. Construindo capacidade: **Relatório Técnico**, v.2, p.307. 2011.
- MELO, V.S.; MEDEIROS, R.M.; SOUSA, F.A.S. **Impactos do clima urbano em Bom Jesus – PI**.
- MELO, V.S.; MEDEIROS, R.M.; SOUZA, L.F.; SOUSA, E.P.; SOUZA, F.A.S. **O clima urbano e seus impactos em Parnaguá – PI**. **Revista Agropecuária Científica no Semiárido**, v.12, n.2, p.181-198, 2016.
- MEYER, P. L. **Probabilidade: aplicações à estatística**. Rio de Janeiro: LTC, 2001.
- MOREIRA, P. S. P. et al. Distribuição e probabilidade de ocorrência das chuvas no município de Nova Maringá-MT. **Revista de Ciências Agroambientais**, Alta Floresta, v. 8, n. 1, p. 9-20, 2010.
- MORETTIN, P. A.; TOLOI, C. M. C. **Séries Temporais**. São Paulo: Atual, 2005. 420 p.
- MURTA, R. M. et al. Precipitação pluvial mensal em níveis de probabilidade pela distribuição gama para duas localidades do sudoeste da Bahia. **Ciência e Agrotecnologia**, Lavras, v. 29, n. 5, p. 988-994, 2005.

- NOBRE, P.; MELO, A. B. C. **Variabilidade climática intrasazonal sobre o Nordeste do Brasil em 1998 – 2000**. Climanálise, CPTEC/INPE, São Paulo. Dezembro, 2001.
- ORTOLANI, A. A.; CAMARGO, M. B. P. **Influência dos fatores climáticos na produção**. Ecofisiologia da Produção Agrícola. Piracicaba: Potafos, 249 p. 1987.
- PAULA, R. K.; BRITO, J. I. B.; BRAGA, C. C. **Utilização da análise de componentes principais para verificação da variabilidade de chuvas em Pernambuco**. XVI Congresso Brasileiro de Meteorologia. Anais... Belém do Pará, PA. CD Rom. 2010.
- REBOUÇAS, Aldo da C.. **Águas doces no mundo e no Brasil**. In: Águas doces do Brasil: Capital ecológico, uso e conservação. Org. REBOUÇAS, Aldo da C. BRAGA, Benedito. TUNDISI, José Galizia. 3º ed. Editora Escrituras. São Paulo. 2006. pág. 01-35.
- SANTANA, M. O.; SEDIYAMA, G. C.; RIBEIRO, A., SILVA, D. D. **Caracterização da estação chuvosa para o estado de Minas Gerais**. Revista Brasileira de Agrometeorologia, v.15, n.1, p.114-120, 2007.
- SANTOS, G. S.; NORI, P. G. OLIVEIRA, L. F. C. Chuvas intensas relacionadas à erosão hídrica. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.14, p.115-123, 2010.
- SILVA, A. M. **Rainfall erosivity map for Brazil**. Catena, v.57, p.251- 259, 2004
- SILVA, J. B. **Análise do Comportamento das Chuvas em Algumas Localidades do Semi-árido do Nordeste Brasileiro** – Revista Brasileira de Agrometeorologia, Santa Maria, V13, n1, p. 173-181, 2005.
- SUDENE. Superintendência do Desenvolvimento do Nordeste - Dados pluviométricos mensais do Nordeste – Série pluviometria 5. Estado da Paraíba. Recife, 239p. 1990.
- TRIOLA, M. F. **Introdução à estatística**. Tradução: Alfredo Alves de Faria. 7. ed. Rio de Janeiro: LTC, 1999. 410 p.
- WALPOLE, R. E. **Probabilidade & estatística para engenharia e ciências**. 8. ed. São Paulo: Pearson Prentice Hall, 20.