Revista Brasileira de Climatologia
ISSN: 2237-8642 (Eletrônica)

ANÁLISE ESTATÍSTICA MULTIVARIADA DA PRECIPITAÇÃO DO ESTADO DE SERGIPE ATRAVÉS DOS FATORES E AGRUPAMENTOS

SANTOS, Eucymara França Nunes - eucymara@gmail.com
Universidade Federal do Sergipe / UFS

SOUSA, Inajá Francisco - inajafrancisco@gmail.com
Universidade Federal do Sergipe / UFS

RESUMO: Neste trabalho utilizou-se as técnicas de análise multivariada, análise fatorial e análise de agrupamentos para sintetizar as informações, e facilitar o entendimento da variabilidade temporal e espacial da precipitação, dos munícipios do Estado de Sergipe no Brasil. Os resultados obtidos mostraram que os meses de março a outubro foram os que apresentaram os mais altos níveis de precipitação, enquanto que os diferentes grupos formados foram encontrados nos munícipios concordantes das regiões litoral, agreste e sertão do Estado. A análise fatorial e análise de agrupamentos apresentaram resultados similares na identificação da estrutura temporal da precipitação em todos os municípios do Estado de Sergipe.

PALAVRAS-CHAVE: precipitação, análise fatorial, análise de agrupamentos.

MULTIVARIATE STATISTICAL ANALYSIS OF THE PRECIPITATION OF THE STATE OF SERGIPE THROUGH FACTORS AND GROUPS

ABSTRACT: In this work the techniques of multivariate analysis, factorial analysis and cluster analysis were used to synthesize the information, and to facilitate the understanding of the temporal and spatial variability of the precipitation of the municipalities of the State of Sergipe in Brazil. The results showed that the months of March to October were the ones that presented the highest levels of precipitation, while the different groups formed were found in the concordant municipalities of the coastal, agreste and sertão regions of the State. Factor analysis and cluster analysis showed similar results in identifying the temporal structure of precipitation in all municipalities of the State of Sergipe.

KEYWORDS precipitation, factor analysis, analysis of clusters.

INTRODUÇÃO

Os aspectos climáticos do semiárido brasileiro resultam em intensas características naturais dessa região principalmente no que diz respeito à disponibilidade de água para uso geral pela população residente. O semiárido nordestino possui em média cerca de 700 mm de chuva por ano, no entanto, essa média poderá variar significativamente dependendo das condições climáticas. Além da relativa baixa disponibilidade de chuvas, a sua distribuição é bastante irregular no tempo e no espaço, o que ocasiona um fator limitante ao desenvolvimento social e econômico da região.

A precipitação pluvial é um dos elementos meteorológicos que exerce grande influência sobre as condições ambientais, sobre a economia, além do efeito direto sobre o balanço hídrico climatológico também exerce grande influência sobre outras variáveis, como a captação de água de chuva no sertão. Tendo em vista o impacto ambiental que a anomalia desse componente provoca no clima regional, verifica-se a importância de se estudar cada vez mais essa

variável considerada fundamental para o desenvolvimento da região Nordeste do Brasil.

O clima exerce grande influência sobre o ambiente, atuando como fator de interações entre componentes bióticos e abióticos. O clima de toda e qualquer região, situada nas mais diversas latitudes do globo, não se apresenta com as mesmas características em cada ano (Soriano, 1997). Apesar de chover tanto quanto em muitas regiões do mundo, a região do Nordeste do Brasil é periodicamente afetada pela ocorrência de secas agrícolas e/ou hidrológicas de elevado grau de severidade, causando perdas parciais ou totais da agropecuária, além de comprometer o abastecimento de água à população. Isso se deve, principalmente, a irregularidade da estação chuvosa na região, caracterizada pela ocorrência de chuvas intensas e de curta duração.

Em estudo realizado por Lundgren, et at. (2015) onde foram aplicadas várias distribuições de probabilidades para compreender o comportamento do índice pluviométrico mensal no Estado de Sergipe, foram utilizados série histórica de chuva para os anos de 1912 a 2012. Os autores concluíram que, as chuvas mensais, não puderam ser ajustadas em todos os meses pelas distribuições de probabilidades testadas. Já Nascimento, et al. (2017), utilizaram técnicas da análise multivariada para encontrar padrões espaciais e temporais do SPI-6, e quantificar os eventos extremos secos para o Estado do Maranhão, de acordo com sua frequência e período para os anos de 1987 a 2015. Os resultados desta pesquisa mostraram que pela escala semestral do SPI, o estado encontra-se dividido em cinco regiões homogêneas, e que o modo de variabilidade do SPI-6 explica bem o regime pluviométrico, e ainda concluíram que as maiores incidências de eventos de secas ocorreram no sul e as maiores chuvas foram registradas no norte do Estado. Em artigo publicado recentemente pelos autores Silva et al. (2017), utilizando dados de diversas variáveis climáticas para o Estado de Sergipe no período de 2011 a 2013, aplicaram a técnica de componentes principais e análise de agrupamento. Os resultados obtidos com esta pesquisa, mostraram que da relação de 10 variáveis climáticas analisadas apenas três: umidade relativa do ar, temperatura do ar máxima e velocidade máxima dos ventos, obtiveram um percentual de explicação de 75,15%.

Na agricultura de sequeiro, a precipitação constitui na variável meteorológica de maior importância para se conseguir ótima produção agrícola. A chuva em determinada época do ano pode ser útil ou prejudicial à agricultura, dependendo se coincide com o período vegetativo ou de colheita de determinadas culturas. Neste trabalho utilizou-se da análise multivariada que por definição representa um conjunto de técnicas estatísticas que processa informações de modo a simplificar a estrutura dos dados, sintetizar as informações e facilitar o entendimento do relacionamento existente entre as variáveis do processo, que são medidas simultaneamente numa mesma unidade experimental, em cada elemento amostral. Como ferramentas para o desenvolvimento dessas metodologias são empregadas técnicas como a análise de componentes principais, análise fatorial, análise de agrupamentos entre outras que permitem a redução de dados ou simplificação estrutural, ordenação e agrupamento e a investigação da dependência entre variáveis (MINGOTI, 2005).

ISSN: 2237-8642 (Eletrônica)

Diante da importância que a precipitação pluviométrica exerce sobre o planejamento agrícola e no desenvolvimento da região Nordeste do Brasil, objetivou-se neste trabalho, avaliar a variabilidade temporal e espacial da precipitação nos munícipios do Estado de Sergipe através de técnicas estatísticas multivariadas de análise fatorial e análise de agrupamentos.

MATERIAL E MÉTODOS

O banco de dados utilizado para as análises deste trabalho foi objeto de coleta no site da Empresa de Desenvolvimento Agropecuário - EMDAGRO, oriundos de 39 municípios do Estado de Sergipe, no período de 1912 a 2012, totalizando 100 anos para as estações mais completas. Para tanto, foram utilizadas as coordenadas geográficas latitude, longitude em graus e altitude em metros, e a precipitação média em séries mensal.

Os dados foram organizados na planilha do Excel e posteriormente importada para o R-Studio, programa gratuito no qual foram realizadas as análises. O primeiro procedimento utilizado foi uma análise estatística descritiva dos dados, utilizando a função summary para obter informações básicas a respeito da principal variável (precipitação) a ser estudada, nos diferentes períodos. Em seguida as análises multivariadas foram obtidas através das funções: matcor, precept, varimax para a análise fatorial, hclust, vegdist e d.cofen para a análise de agrupamentos.

Os valores médios que representam o período com maior intensidade de precipitação estão entre os meses de abril a agosto, oscilando entre 102,50 mm a 175,20 mm, mostrados na Tabela 1, intervalo onde se encontra o valor médio do período chuvoso, os meses de março e setembro também pertencem a este período, os quais não apresentam uma média alta de precipitação, mas ainda foram um pouco maiores comparando com os meses de outubro, novembro, dezembro, janeiro e fevereiro referentes ao período seco.

Tabela 1 - Estatística descritiva da variável resposta precipitação por períodos, mensal, anual, chuvoso e seco, dos munícipios de Sergipe, com dados de 1912 a 2012.

Período	Média (mm)	D. P.	Med. (mm)	Mín. (mm)	Máx. (mm)
Janeiro	46,73	9,36	48,90	29,10	68,30
Fevereiro	56,35	18,35	53,50	21,40	96,10
Março	82,48	20,81	78,60	48,80	128,20
Abril	134,23	47,06	118,70	61,40	216,30
Maio	175,20	67,17	164,60	63,80	349,60
Junho	163,70	51,74	154,50	63,20	266,10
Julho	150,90	46,07	142,70	55,30	229,20
Agosto	102,55	34,74	104,90	31,20	168,50
Setembro	63,42	23,58	65,40	16,30	100,94
Outubro	43,97	16,18	46,60	12,10	72,29
Novembro	43,72	13,55	44,60	19,20	70,60
Dezembro	39,20	7,47	38,80	24,00	54,10
Anual	92,63	27,23	91,20	41,80	137,70

D.P.: desvio padrão; Med.: mediana (mm); Mín: valor mínimo da série (mm); Máx: valor máximo da série (mm).

ISSN: 2237-8642 (Eletrônica)

A Figura 1 (A) mostra os pontos das estações com suas respectivas localizações no mapa do Estado de Sergipe, as 39 estações pluviométricas seguindo as latitudes e longitudes, mostrando a dispersão de pontos no território sergipano. A Figura 1 (B) evidencia as regiões agreste, litoral e sertão no Estado de Sergipe.

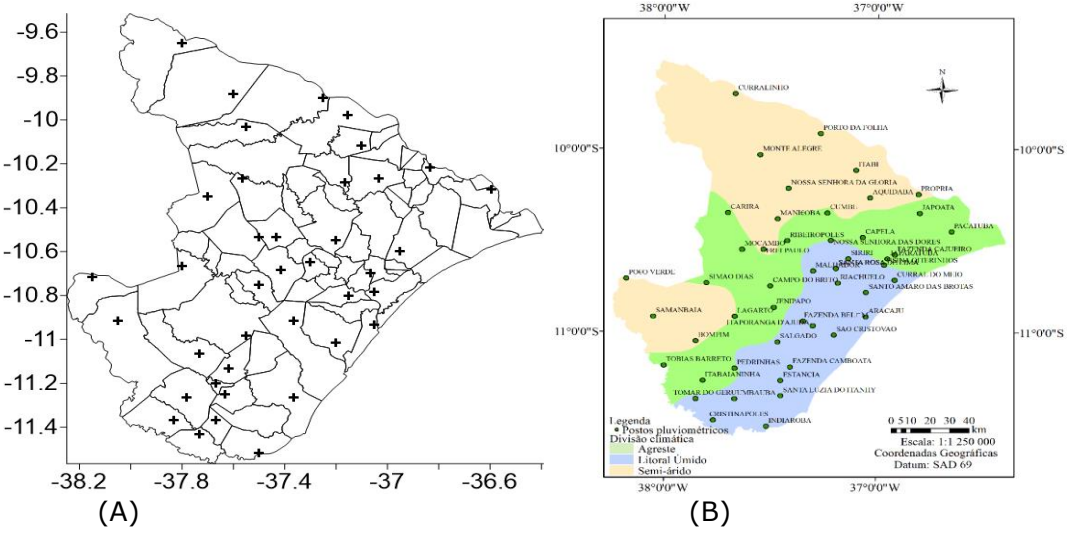


Figura 1 – (A) Espacialização das estações analisadas no estado de Sergipe (B) Mapa do Estado de Sergipe.

ANÁLISE FATORIAL

A análise fatorial é uma técnica de análise multivariada com o objetivo de avaliar a estrutura das variáveis através das suas correlações, sendo que as variáveis dentro de um grupo particular sejam altamente correlacionadas entre si, e mantenham o mínimo de correlação com variáveis em grupos diferentes, então cada grupo de variáveis representa um único fator que é responsável pelas correlações observadas (JOHNSON e WICHERN, 1998).

Os fatores explicam a variância das variáveis observadas, tal como se revelam pelas correlações das variáveis que estão sendo analisadas. O método mais conhecido para a extração dos fatores é feito por meio da análise de componentes principais, que é baseado no pressuposto que podem definir vetores estatisticamente não correlacionados a partir de combinações lineares das variáveis originais. Segundo Mingoti (2005), o modelo é representado da seguinte forma:

ISSN: 2237-8642 (Eletrônica)

$$\begin{aligned} Z_1 &= l_{11}F_1 + l_{12}F_2 + \ldots + l_{1m}F_m + \varepsilon_1 \\ Z_2 &= l_{21}F_1 + l_{22}F_2 + \ldots + l_{2m}F_m + \varepsilon_2 \\ &\vdots \\ Z_p &= l_{p1}F_1 + l_{p2}F_2 + \ldots + l_{pm}F_m + \varepsilon_p \end{aligned} \quad (1)$$

onde F_{mx1} é um vetor aleatório que contém m fatores chamados de variáveis latentes, relacionadas linearmente com as variáveis originais, o vetor ε representa os erros aleatórios, que não são explicados pelos fatores comuns, o coeficiente l_{ij} são os pesos correspondentes ao grau de relacionamento linear entre Zi e Fi.

As variáveis são definidas como combinações lineares dos fatores comuns que irão explicar a parcela da variância de cada variável, mais um desvio que resume a parcela da variância total não explicada por estes fatores. A parcela explicada pelos fatores é chamada de comunalidade, e a parcela não explicada de especificidade (HAIR el al, 2005).

A quantidade de fatores extraídos seguindo o critério de Kaiser, que recomenda desprezar aqueles fatores cujos autovalores (variância explicada) sejam menores do que um, outro critério conhecido é o do scree plot, na representação gráfica do autovalor versus o percentual da variância explicada, desconsidera-se aqueles em que estejam praticamente paralelos ao eixo das componentes (REIS, 2001).

A interpretação depende tanto da experiência do pesquisador quanto do poder da explicação das variáveis. A primeira solução geralmente fornece um fator bem correlacionado com muitas variáveis, fornecendo informações que não são fáceis de interpretar. O efeito final de rotacionar a matriz é redistribuir a variância dos fatores, para podermos realizar uma interpretação mais adequada. Segundo Hair et. al. (2005), o método de rotação ortogonal mais comumente utilizado é o VARIMAX.

ANÁLISE DE AGRUPAMENTOS

A análise de agrupamentos tem como objetivo agrupar as variáveis, conforme a sua proximidade e características comuns, buscando mostrar a homogeneidade dentro dos grupos e heterogeneidade entre os grupos. Consistem em transformar um conjunto de variáveis originais em um pequeno número de combinações lineares, os chamados componentes principais, de dimensões equivalentes, e com propriedades importantes.

Os escores (não correlacionados) dos componentes, colocados uns contra os outros, podem favorecer a análise de gráficos, onde se pode observar a formação de grupos, a representação gráfica dos componentes principais são ferramentas valiosas na explanação na análise de dados. O objetivo é obter variáveis que retenham o máximo possível de informações nelas contidas e

ISSN: 2237-8642 (Eletrônica)

expliquem a maior parte da variabilidade total, revelando que tipo de relacionamento existe entre eles.

Um conceito fundamental é a escolha de um critério que meça a distância entre dois objetos, ou que quantifique o quanto eles são parecidos, esta medida é chamada coeficiente de dissimilaridade (quanto maior o valor, maior a diferença entre os objetos). A medida mais comum e apropriada para variáveis quantitativas e a mais utilizada é: (Bussab, 1990).

1) Distância Euclidiana – a distância entre dois elementos (xi e xj) é a raiz quadrada do somatório dos quadrados das diferenças entre valores de i e j para todas as variáveis (v = 1, 2, ..., p).

$$d(x_i, x_j) = \sqrt{\sum_{v=1}^{p} (x_{iv} - x_{jv})^2} \quad (2)$$

Os métodos de agrupamento têm o objetivo de reunir entidades em grupos homogêneos. O método de classificação denominado hierárquico consiste em reunir indivíduos em grupos, e o processo repete-se em diferentes níveis até formar uma árvore chamada dendograma. Segundo Chatfield (1997), o tipo de algoritmo é chamado de aglomerativo porque ele opera por uma série de união, começando com n grupos de apenas um indivíduo e termina em um grupo de n indivíduos, os mais utilizados são:

1) Ligação simples – esse método tem seu procedimento iniciado com a procura dos dois objetos mais similares na matriz de dissimilaridade. Ele consiste em reconhecer os indivíduos mais próximos os quais são reunidos formando o grupo inicial, calcula-se então a distância daquele grupo em relação aos demais indivíduos e nos estágios mais avançados. A distância entre os progenitores k e um grupo formado pelos progenitores i e j é dada por:

$$d_{(ij)k} = \min (d_{ik}; d_{jk}) \quad (3)$$

2) Ligação completa – este método, após agrupar os dois indivíduos mais semelhantes de menor distância, verifica a distância máxima deste primeiro grupo para os objetos restantes, Frei (2006). As distâncias entre um grupo e um indivíduo devem ser calculadas pela expressão:

$$d_{(ij)k} = \text{máx} (d_{ik}; d_{jk}) \quad (4)$$

3) Ligação média – este método trata a distância entre dois conglomerados como a média das distâncias entre todos os pares de elementos que podem ser formados com os elementos dos dois conglomerados que

estão sendo comparados. As distâncias entre o objeto que se quer incluir num grupo são:

$$d(C_1, C_2) = \sum_{i \in C_1} \sum_{j \in C_2} \left(\frac{1}{n_i n_j} \right) d(x_i, x_j) \qquad (5)$$

onde C1 e C2 são grupos aos quais pertencem os elementos ni e nj, e estes são os números de elementos nos grupamentos i e j.

4) Centróide – este método define a coordenada de cada grupo como sendo a média das coordenadas de seus objetos. A distância entre os grupos é obtida através do cálculo das distâncias entre os centróides.
5) Ward – neste método um grupo será reunido a um outro se essa união proporcionar o aumento da variância intragrupo, Valentin (2000). A variância intragrupo será calculada para todas as alternativas de aglomeração, escolhendo a que proporcionará a menor variância, sendo aplicado a todos os passos da análise. A distância é definida como:

$$d(x_i, x_j) = \left(\frac{n_i n_j}{n_i + n_j} \right) (\bar{x}_i - \bar{x}_j)' (\bar{x}_i - \bar{x}_j) \qquad (6)$$

onde ni e nj são os números de elementos nos grupamentos i e j.

RESULTADOS E DISCUSSÃO

A dispersão da precipitação na Figura 2 mostra a variabilidade dos totais médios da precipitação nas 39 estações localizadas no Estado de Sergipe. Os maiores valores se concentram no litoral e em algumas áreas no nordeste do agreste sergipano.

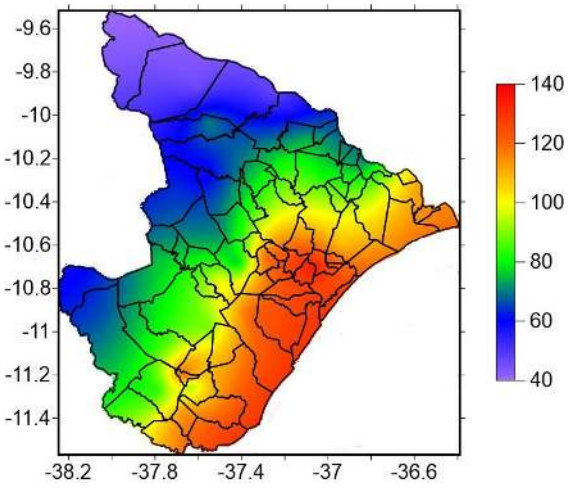


Figura 2 – Distribuição da precipitação média anual no Estado de Sergipe

ISSN: 2237-8642 (Eletrônica)

A Tabela 2 revela a intensidade do relacionamento entre os meses do ano da precipitação, medida mais adequada para avaliar o grau de relacionamento entre as variáveis, assim a matriz de correlação foi utilizada para os cálculos da análise fatorial, podemos observar altas correlações entre os meses de março a outubro, janeiro possui correlação moderada com os meses, novembro tem um bom relacionamento com fevereiro, março e outubro, e dezembro não possui correlação significativa com os demais meses.

Tabela 2 – Matriz de Correlação da Precipitação Mensal

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov
Fev	0,66										
Mar	0,66	0,82									
Abr	0,66	0,74	0,88								
Mai	0,55	0,62	0,80	0,93							
Jun	0,61	0,64	0,80	0,94	0,86						
Jul	0,59	0,62	0,79	0,92	0,88	0,97					
Ago	0,56	0,72	0,81	0,88	0,84	0,94	0,96				
Set	0,58	0,76	0,79	0,90	0,85	0,90	0,90	0,94			
Out	0,58	0,80	0,87	0,91	0,86	0,86	0,85	0,89	0,94		
Nov	0,55	0,80	0,81	0,72	0,68	0,66	0,67	0,79	0,79	0,86	
Dez	0,28	0,52	0,51	0,26	0,13	0,19	0,19	0,30	0,32	0,42	0,63

A análise fatorial foi realizada através do método das componentes principais utilizando a matriz de correlação, foi observado que 86,45% do total da variância pode ser explicada por dois fatores (Tabela 3), com a análise dos autovalores maiores que 1,0, seguindo o critério de Kaiser. O mesmo resultado foi obtido com critério do Scree Plot apresentado na Figura 3, o gráfico sugere que a partir do segundo componente, a variância dos componentes torna-se praticamente nula. A razão para isto é que as variáveis correlacionadas aos componentes principais explicam menores percentuais de variância.

Tabela 3 - Total da variância explicada

	Componentes Principais		
Componente	Autovalores	Variância (%)	Acumulada (%)
1	9,0671	75,56	75,56
2	1,3073	10,89	86,45
3	0,6111	5,09	91,55
4	0,3004	2,50	94,05
5	0,2470	2,06	96,11
6	0,1847	1,54	97,65
7	0,1143	0,95	98,60
8	0,0673	0,56	99,16
9	0,0463	0,39	99,55
10	0,0274	0,23	99,78
11	0,0179	0,15	99,93
12	0,0090	0,07	100,00

ISSN: 2237-8642 (Eletrônica)

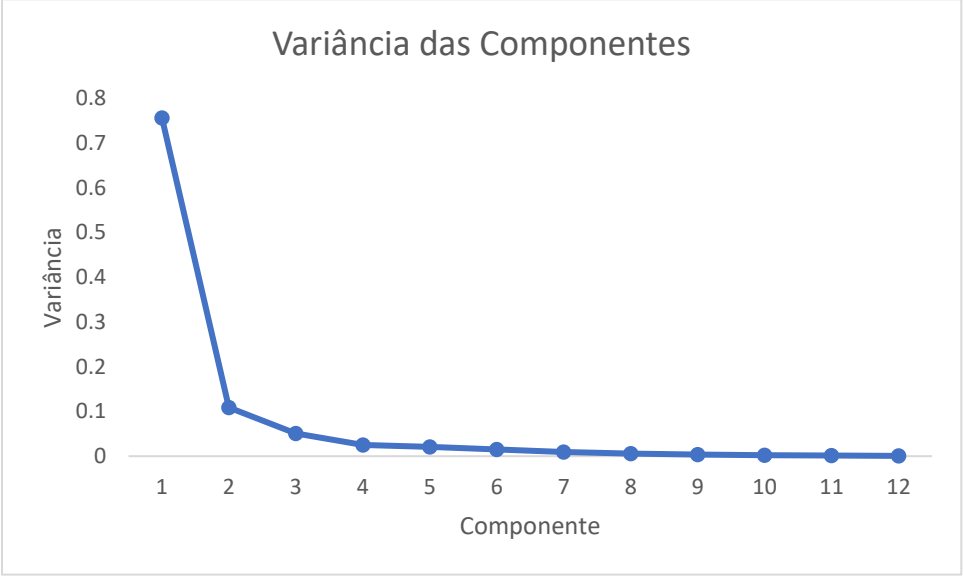


Figura 3 – Scree Plot das componentes principais

Um dos fatores que indica a aplicação da análise fatorial é a medida de Kaiser-Meyer-Oklin (KMO), índice que avalia a adequação da análise aos dados, este valor pode variar entre 0 e 1, onde 0,8 ou acima é aceitável, para o caso aqui estudado o valor apresentado foi de 0,8690, onde verifica-se a utilização da técnica aos dados.

No trabalho de Resende et al. (2007) foi utilizado a análise fatorial com o objetivo de definir os potenciais de desenvolvimento dos munícipios da região Sul, no Estado de Minas Gerais, a técnica possibilitou a extração de três fatores capazes de explicar 64,5% da variância total, e apresentou um índice de KMO igual a 0,8147.

A partir da extração dos fatores através das componentes principais, o primeiro fator sozinho explica 75,6% da variabilidade dos dados, e o segundo 10,9%, valores apresentados na Tabela 4. A interpretação dos fatores é feita por meio das cargas fatoriais que expressam as covariâncias entre cada fator e as variáveis originais, a identificação de estrutura de dependência não estão sendo bem interpretáveis, praticamente todos os meses estão concentrados no primeiro fator, com exceção dos meses de janeiro e dezembro, mostrado na Tabela 5.

A primeira solução é chamada de não rotacionada, para obtermos uma interpretação mais adequada dos fatores aplicamos o método de rotação ortogonal Varimax, cujos resultados podem ser vistos na Tabela 5. Então houve uma distribuição de contribuição nos fatores, 61,9% no primeiro e 24,6 no segundo fator (Tabela 4).

Guedes et al. (2010) consideraram duas componentes principais rotacionadas, que explicaram 88% da variância total dos dados, facilitando a correlação no tempo e espaço da origem da chuva no estado do Piauí.

ISSN: 2237-8642 (Eletrônica)

A compreensão dos meses nos fatores (Tabela 5) mostrou que de março a outubro a correlação é alta no primeiro fator, corroborando com o resultado visto na matriz de correlação em que os meses de março, maio, junho, agosto e outubro são bem correlacionados com os demais messes, exceto janeiro, fevereiro, novembro e dezembro, dos quais novembro e dezembro estão bem interpretados no segundo fator, o que não ocorre com janeiro e fevereiro. De acordo com (Uvo, 1989), a quadra chuvosa nessa região concentra-se entre os meses de fevereiro a maio, sendo basicamente dependente do deslocamento para o Hemisfério Sul da Zona de Convergência Intertropical (ZCIT), que contribui com aproximadamente de 70 a 80% da distribuição da chuva.

Enquanto que nos setores, Agreste e Litoral do Estado observa-se uma melhor distribuição das chuvas, isso provavelmente ocorre em função do posicionamento geográfico, os valores climatológicos observados chegam a ultrapassar os 1.000 mm. Nesses setores o comportamento pluviométrico é favorecido pela incursão de diversos sistemas meteorológicos na região Leste do NEB, tais como: frentes frias (Kousky, 1979), ondas de leste (Yamazaki e Rao 1977; Chan, 1990), vórtices ciclônicos da troposfera superior (Kousky e Gan, 1981) e sistemas de brisas marítimas-terrestre (Kousky, 1980).

Tabela 4 - Variâncias explicadas das cargas fatoriais antes e depois da rotação VARIMAX.

Cargas			Cargas Rotacionadas		
Total	Variância (%)	Acumulada (%)	Total	Variância (%)	Acumulada (%)
9,07	75,6	75,6	7,43	61,9	61,9
1,31	10,9	86,5	2,95	24,6	86,5

Tabela 5 – Cargas Fatoriais

	Cargas Fatoriais		Cargas Fatoriais Rotacionadas	
Componente	Fator 1	Fator 2	Fator 1	Fator 2
Janeiro	0,6945	0,0827	0,5789	0,3925
Fevereiro	0,8329	0,3291	0,5886	0,6750
Março	0,9226	0,1647	0,7438	0,5702
Abril	0,9566	-0,1719	0,9287	0,2867
Maio	0,8927	-0,2890	0,9257	0,1533
Junho	0,9262	-0,2838	0,9531	0,1733
Julho	0,9226	-0,2990	0,9569	0,1582
Agosto	0,9454	-0,1409	0,9044	0,3091
Setembro	0,9472	-0,0990	0,8868	0,3472
Outubro	0,9599	0,0380	0,8351	0,4748
Novembro	0,8561	0,3701	0,5903	0,7221
Dezembro	0,4157	0,8448	-0,0188	0,9413

O gráfico da Figura 3 evidencia os meses por fator mostrado anteriormente, de março a outubro acima de 0,5 no fator 1, novembro e dezembro no fator 2, contudo janeiro e fevereiro são dois meses bem indefinidos.

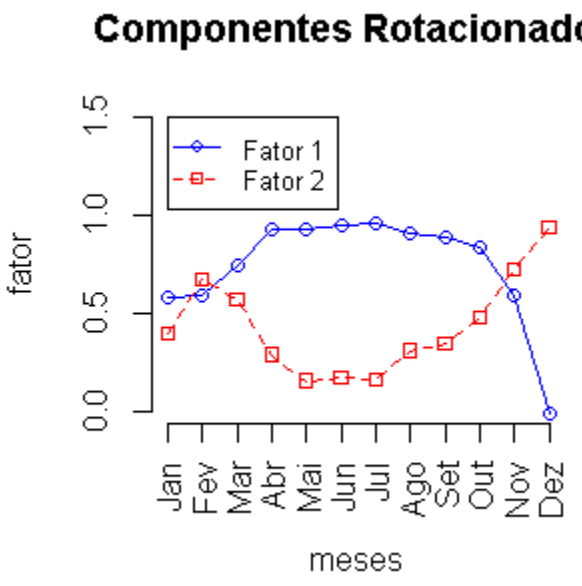
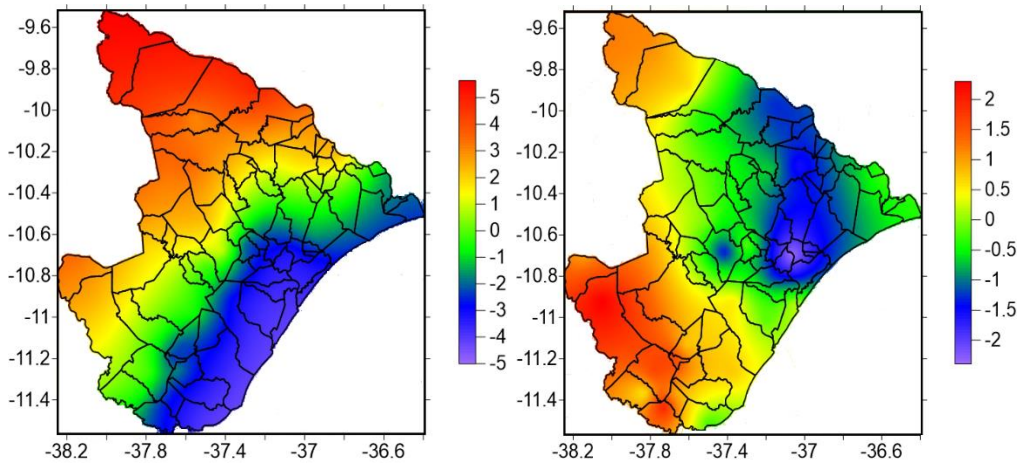


Figura 3 – Gráfico Fator 1 x Fator 2

A Figura 4 mostra os mapas com os escores dos resultados extraídos da análise fatorial, a figura (A) apresenta os maiores valores no norte e noroeste de março a outubro e a figura (B) os maiores pesos estão concentrados no sul e sudoeste em novembro e dezembro no Estado de Sergipe.



(A) Fator 1 (B) Fator 2

Figura 4 – Distribuição espacial dos fatores.

A análise de agrupamentos realizada pela distância euclidiana e os cinco métodos de agrupamentos serão avaliados pelo número, pela homogeneidade

ISSN: 2237-8642 (Eletrônica)

dos grupos e através da correlação cofenética (coeficiente R de Pearson), medida de validação para o grau de ajuste entre a matriz original e a matriz de agrupamentos obtida após a construção do dendograma. Considera-se aceitável um coeficiente cofenético superior a 0,7, sendo que o maior valor encontrado na análise é o critério de escolha do melhor método de agrupamento. Os métodos mais apropriados foram o da ligação média com 0,7677 e o centróide com 0,7617, mas por uma melhor interpretação dos grupos o mais indicado foi o centróide que possibilitou a formação de 7 grupos de estações (Figura 5).

Guedes et al. (2010) utilizaram a distância euclidiana e o método de agrupamento hierárquico de Ward para identificar quatro grupos de precipitação homogênea de 27 estações pluviométricas do estado do Piauí.

Dez variáveis climáticas dos municípios Sergipanos foram avaliadas por Silva et al. (2017), no período de 2011 a 2013. A umidade relativa do ar, a temperatura do ar máxima e a velocidade dos ventos máxima foram suficientes para sumarizar 75,15% das informações, resultante da análise de componentes principais. Com estas três variáveis foi possível obter três grupos de municípios, o primeiro formado por Estância, o segundo Itabaiana e Japaratuba, e o terceiro por Nossa Senhora da Glória, Riachão do Dantas e Poço Redondo. Os municípios Sergipanos agrupados neste trabalho através dos meses de precipitação, com 100 anos de informações, apenas coincidiram no mesmo grupo Poço Redondo e Nossa Senhora da Glória, já Estância e Japaratuba pertencem a um mesmo grupo, Riachão do Dantas e Itabaiana a outro .

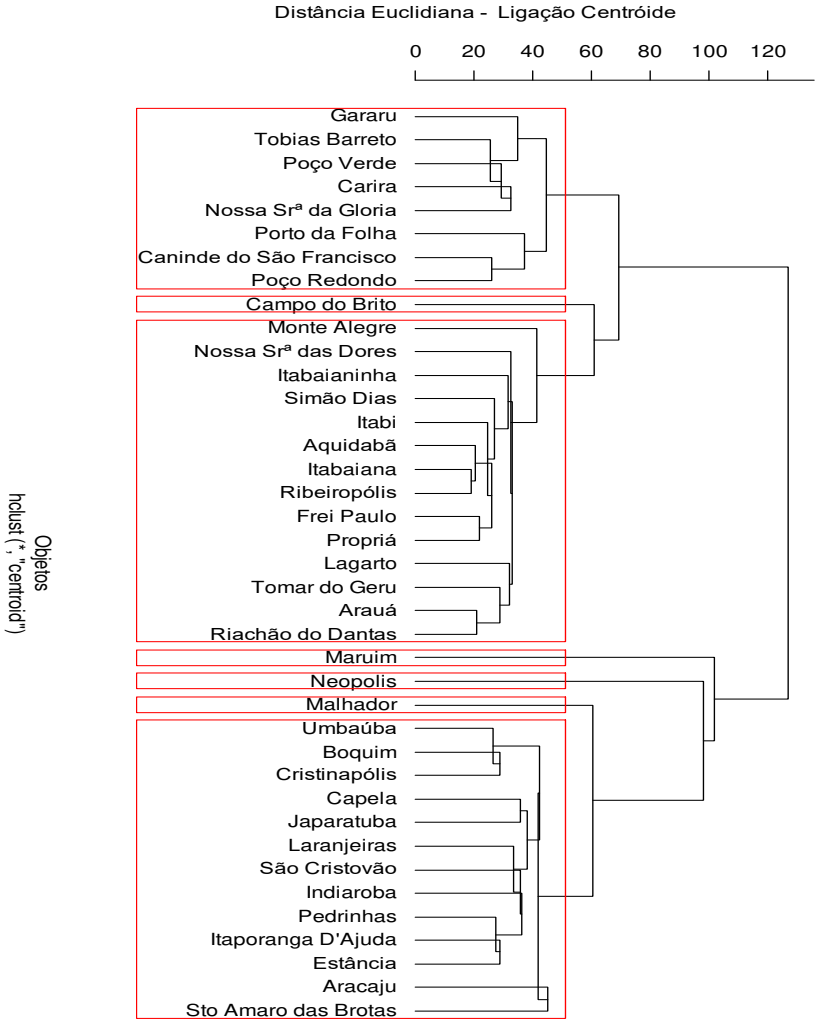
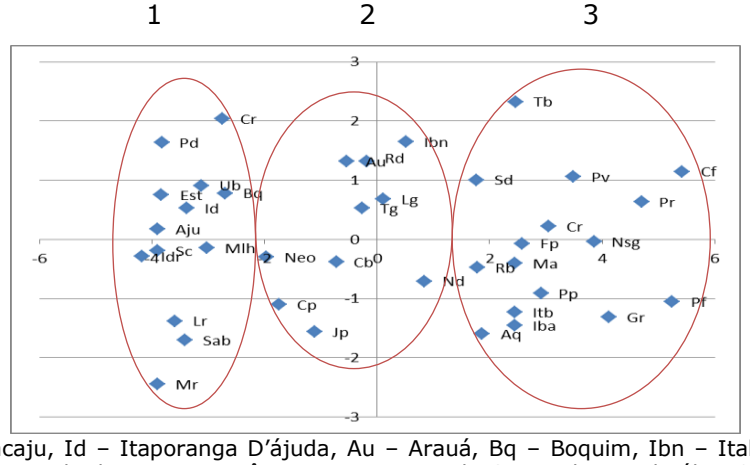


Figura 5 – Dendograma resultante da distância euclidiana e do método agrupamento centroide das 39 estações meteorológicas do Estado de Sergipe.

A Figura 6 mostra os escores ou pesos dos fatores nas estações para visualizar os três grupos formados, identificando municípios do litoral no primeiro grupo, os que estavam isolados se alocaram em outros, o segundo grupo apresenta regiões do litoral e do agreste e o terceiro predominantemente do sertão, mostrando a semelhança com os grupos formados pelo método de agrupamentos.

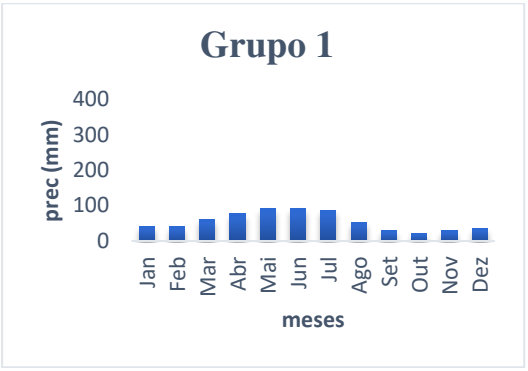


Nota: Aju – Aracaju, Id – Itaporanga D'ájuda, Au – Arauá, Bq – Boquim, Ibn – Itabaianinha, Cr – Cristinápolis, Pd – Pedrinhas, Est – Estância, Tg – Tomar do Geru, Ub – Umbaúba, Cr – Carira, Lg – Lagarto, Rd – Riachão do Dantas, Sd – Simão Dias, Pv – Poço Verde, Tb – Tobias Barreto, Cf – Canindé do São Francisco, Pr – Poço Redondo, Sab – Santo Amaro das Brotas, Sc – São Cristóvão, Lr – Laranjeiras, Mr – Maruim, Idr – Indiaroba, Cp – Capela, Jp – Japaratuba, Mlh – Malhador, Cb – Campo do Brito, Iba – Itabaiana, Rb – Ribeiropólis, Fp – Frei Paulo, Nd – Nossa Senhora das Dores, Aq – Aquidabã, Itb – Itabi, Gr – Gararu, Ma – Monte Alegre, Nsg – Nossa Senhora da Glória, Pf – Porto da Folha, Pp – Propriá, Neo - Neópolis

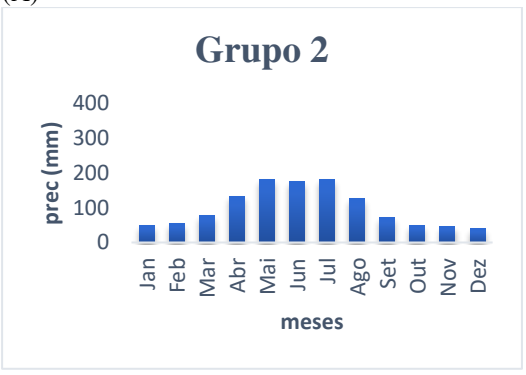
Figura 6 – Distribuição dos grupos pelos escores das cargas fatoriais dos 39 munícipios.

Na Figura 8 os meses de março a agosto apresentam maior nível de precipitação em todos os grupos, verificamos a semelhança na precipitação mensal entre os grupos 1, 2 e 3 predominantes de munícipios do sertão, observamos nos grupos 4 e 5 uma sinuosa diferença entre os meses de maio a junho. Os grupos 6 e 7 formados por maioria de munícipios do litoral e alguns do agreste, são bem parecidos, com maiores precipitações em julho e maio respectivamente.

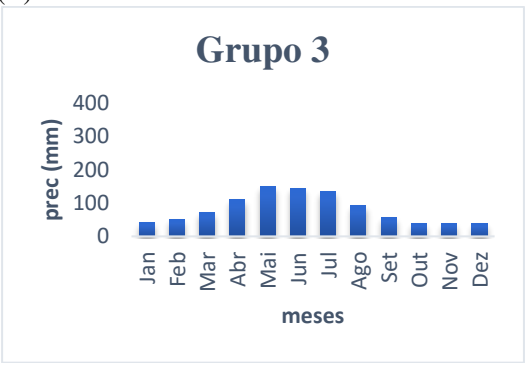
ISSN: 2237-8642 (Eletrônica)



(A)



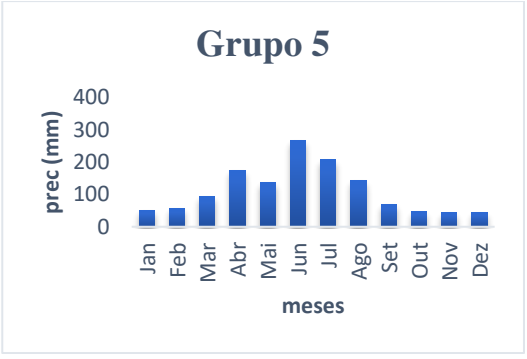
(B)



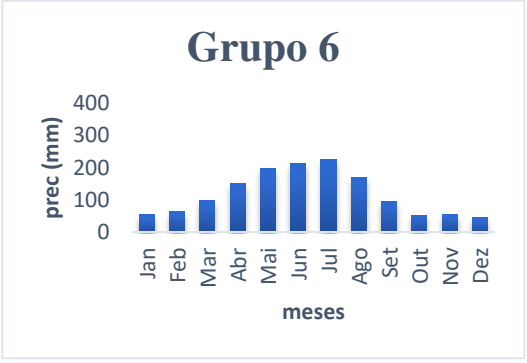
(C)



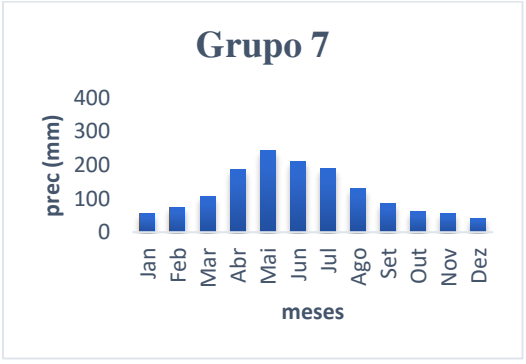
(D)



(E)



(F)



(G)

Figura 4 – Precipitação média anual dos grupos formados pelo método de agrupamentos. (A) Carira, Nossa Senhora da Glória, Poço Verde, Tobias Barreto, Gararu, Canindé do São Francisco, Poço Redondo, Porto da Folha. (B) Campo do Brito, (C) Arauá, Riachão do Dantas, Tomar do Geru, Lagarto, Nossa Senhora das Dores, Itabaianinha, Itabaiana, Ribeirópolis, Aquidabã, Itabi, Frei Paulo, Propriá, Simão Dia, Monte Alegre, (D) Maruim, (E) Neópolis, (F) Malhador, (G) Aracaju, Santo Amaro das Brotas, São Cristóvão, Laranjeiras, Indiaroba, Itaporanga, Estância, Pedrinhas, Capela, Japaratuba, Cristinápolis, Umbaúba, Boquim.

CONCLUSÃO

A análise fatorial e análise de agrupamentos apresentaram resultados similares na identificação da estrutura temporal da precipitação nas cidades do Estado de Sergipe. As técnicas estatísticas apresentam grupos concordantes dos munícipios em relação à precipitação e a região, identificou os meses de maior nível de precipitação de março a outubro. Os grupos dos munícipios foram corretamente separados nas três regiões do Estado, formou-se um grupo predominantemente pertencente ao sertão sergipano, outro grupo com municípios localizados na região litorânea e por último os municípios localizado na região agreste e alguns municípios que fazem fronteira com a região do litoral.

REFERÊNCIAS

BUSSAB, W. DE O.; MIAZAKI, E. S.; ANDRADE, D. F. DE. Introdução à análise de agrupamentos, São Paulo. Associação Brasileira de estatística. 9º Simpósio Nacional de Probabilidade e Estatística, 1990. 105 p.

CHAN, C.S., 1990: Análise de distúrbios ondulatórios de leste sobre o Oceano Atlântico Equatorial Sul. (INPE-5222-TDL/437), 134p.

CHATFIELD, C.; COLLINS, A. J. Introduction to multivariate analysis. London. Chapman e Hall, 1997.

EMPRESA DE DESENVOLVIMENTO AGROPECUÁRIO DE SERGIPE. Disponível em < http://www.emdagro.se.gov.br/> Acesso em: 24 Agosto de 2015.

FREI, F. Introdução à análise de agrupamentos. São Paulo: Editora UNESP, 2006. 111 p.

GUEDES, R. V. DE. S.; LIMA, F. J. L. DE.; AMANAJÁS, J. C.; BRAGA, C. C. Análise em componentes principais da precipitação pluvial no Estado do Piauí e agrupamento pelo método de ward. Revista de Geografia, v. 27, n. 1, jan/mar. 2010.

HAIR JR.; ANDERSON, R.; TATHAM, R.; BLACK, W. Análise multivariada de dados. 5 ed. Porto Alegre: Artmed, 2005. 593 p.

JOHNSON, R. A.; WICHERN, D. W. Applied multivariate statistical analysis. 4 ed. New Jersey: Pratice Hall, 1998. 816 p.

KOUSKY, V. E., 1979, "Frontal influences on Northeast Brazil", Mon. Wea, Rev., 107, 1140-1153.

KOUSKY, V. E., e GAN, A., 1981, "Upper tropospheric cyclonic vortices in the tropical South Atlantic", Tellus, 33, 538-551.

KOUSKY, V. E. 1980, "Diurnal rainfall variation on Northeast Brazil" Mon. Wea. Rew 108, 488-498.

KOUSKY, V. E., e GAN, A., 1981, "Upper tropospheric cyclonic vortices in the tropical South Atlantic", Tellus, 33, 538-551.

LUNDGREN, W. J. C., SOUSA, I. F. de., NETTO. A. O. Uso de Distribuições de Probabilidades para Ajuste aos Dados de Precipitação Mensal do Estado de Sergipe. Revista Brasileira de Geografia Física v. 08 n. 01. 2015, p. 071-080.

NASCIMENTO, F.C.A., BRAGA, C. C., ARAÚJO, F. R. C. D. Análise Estatística dos Eventos Secos e Chuvosos de Precipitação do Estado do Maranhão. Revista Brasileira de Meteorologia. v. 32, n. 3, 375-386, 2017.

MINGOTTI, S. A. Análise de dados através de métodos de estatística multivariada. Belo Horizonte: Editora UFMG, 2005, 295 p.

REIS, E. Estatística multivariada aplicada. 2 ed. Lisboa: Sílabo, 2001. 253 p.

RESENDE, M. L.; FERNANDES, L. P. DE S.; SILVA, A. M. R. E. Utilização da análise fatorial para determinar o potencial de crescimento econômico em uma região do sudeste do Brasil. Revista Economia e Desenvolvimento, n. 19, 2007.

SILVA, C. M. A. DA.; BARRETO, I. D. DE. C.; SANTOS, E. A. B. DOS.; BORGES, P. DE. F.; ARAÚJO, L. S. DE.; SANTOS, E. F. N. Análise das variáveis climáticas das estações agrometereológicas do estado de Sergipe através de métodos multivariados (2011-2013). Revista Gaia. V. 11(1): 144-156, 2017.

SORIANO, B. M. A. Caracterização climática de Corumbá -MS. Corumbá: EMBRAPA-CPAP, 1997. 25p. (EMBRAPA-CPAP. Boletim de Pesquisa, 11).

UVO, C. B. A Zona de Convergência Intertropical (ZCIT) e sua relação com a precipitação na Região Nordeste do Brasil. 1989. 99 f. Dissertação (Mestrado em Meteorologia) – INPE, São José dos Campos, 1989.

VALENTIN, J. L. Ecologia numérica: Uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Editora Interciência, 2000. 117 p.

ISSN: 2237-8642 (Eletrônica)

YAMAZAKI, Y.; RAO, V. B. Tropical cloudness over the south Atlantic Ocean. Journal of the Meteorological Society of Japan. V. 55, 203-207, 1977.