

# ANÁLISE EM COMPONENTES PRINCIPAIS DA PRECIPITAÇÃO PLUVIAL NO ESTADO DO PIAUÍ E AGRUPAMENTO PELO MÉTODO DE WARD

Roni Valter de Souza GUEDES<sup>1</sup>

Francisco José Lopes de LIMA<sup>1</sup>

Jonathan Castro AMANAJÁS<sup>1</sup>

Célia Campos BRAGA<sup>2</sup>

## RESUMO

A precipitação pluvial sofre variação espacial e temporal, sendo dependente de inúmeros fatores geográficos e atmosféricos locais e de grande escala. O aumento das séries históricas das variáveis necessárias a previsão e estudos do tempo e clima, indica a necessidade de estudos com métodos que possam extrair amostras que possuam a maior parte das informações estatísticas da base de dados. Usando o método da Análise em Componentes Principais (ACP), algumas características do padrão pluviométrico mensal de 27 estações no estado do Piauí, foram estabelecidas. Foram consideradas as duas primeiras componentes principais (CPs) rotacionadas, que explicam 88% da variância total dos dados, facilitando a correlação no tempo e espaço da origem da chuva no estado. Foi efetuado também a análise de agrupamento hierárquico utilizando o método de Ward, que identificou quatro grupos de precipitação homogênea e fez-se uma comparação com os setores territoriais explicados pelas CPs encontradas. A CP1 com maior parte da variância, explicou a precipitação de janeiro a agosto, com melhor ajuste para o grupo 1, norte do estado onde se observam os maiores valores de precipitação. Já a CP2 se ajustou bem ao grupo 4, no sul do estado, explicando melhor os meses de setembro a dezembro.

**Palavras-chave:** Precipitação Pluvial; Componentes Principais; Agrupamento.

## ABSTRACT

Rainfall undergoes spatial and temporal variation, being dependent on numerous factors geographic and atmospheric locations and a large scale. The increase in the time series of variables necessary to predict and study the weather and climate, indicates the need for studies with methods that can extract samples that have the most statistical information from the database. Using the Principal Component Analysis (PCA) method, some characteristics of the rainfall pattern monthly of 27 stations in the state of Piauí, are established. Are considered the first two principal components (PCs) rotated, which explain 88% of the total data variance, facilitating the correlation in time and space to the origin of rain in the state. Was also performed hierarchical cluster analysis using Ward's method, which identified four

---

<sup>1</sup> Pós-graduandos do Departamento de Ciências Atmosféricas, UFCG. Av. Aprígio Veloso, 882, Bairro Universitário. Campina Grande – PB, Brasil. CEP: 58.429-140. E-mail: [jonathan.amanajas@hotmail.com](mailto:jonathan.amanajas@hotmail.com); [thescolopes@yahoo.com.br](mailto:thescolopes@yahoo.com.br); [roniguedes84@yahoo.com.br](mailto:roniguedes84@yahoo.com.br)

<sup>2</sup> Professora do Departamento de Ciências Atmosféricas, UFCG. Av. Aprígio Veloso, 882, Bairro Universitário. Campina Grande – PB, Brasil. CEP: 58.429-140. E-mail: [celia@dca.ufcg.edu.br](mailto:celia@dca.ufcg.edu.br)

homogeneous precipitation groups and made a comparison with the territorial sectors explained by the CPs found. The CP1 with most of the variance explained rainfall from January to August, with the best fit for group 1, north of the state where they observe the highest values of precipitation. The CP2 adjusted well to group 4, in the southern state, explaining better the months September to December.

**Key words:** Rainfall; Principal Components; Clustering.

## 1. INTRODUÇÃO

No Estado do Piauí são identificadas três unidades de relevo: Baixada litorânea, ocorre ao norte do território; Planalto de Chapada e Serras, que apresentam os pontos mais elevados do Estado (entre 600 e 880 metros), encontrados a leste, sudoeste e sul; e a planície do Parnaíba <<http://www.brasilecola.com/brasil/aspectos-naturais-piaui.htm>>.

Em decorrência de sua posição, o Piauí caracteriza-se, como uma típica zona de transição, apresentando, conjuntamente, aspectos do semi-árido nordestino, da pré-Amazônia e do Planalto Central do Brasil, refletindo as condições de umidade das diversas zonas. Apresenta caatinga arbórea e arbustiva, cerrado e cerradão, floresta decidual e formações pioneiras de restinga, mangue e aluvial campestre. Este Estado conta com o Rio Parnaíba e alguns de seus afluentes responsáveis pela geração de energia elétrica, através da hidroelétrica de Boa Esperança. No Piauí são identificadas duas características climáticas: tropical quente e úmido, e semi-árido. Apresenta temperaturas médias elevadas, variando entre 18° (mínimas) e 40° C (máximas). A umidade relativa do ar oscila entre 60 e 84%. A frequência de chuvas diminui a medida que se avança para a região sudeste do Estado; porém, níveis anuais médios de precipitação abaixo de 800 mm são encontrados apenas em 35% do território piauiense. O clima semi-árido abrange, principalmente, o sudoeste do Estado. As chuvas são distribuídas nessas regiões de maneira irregular <<http://www.pi.gov.br/piaui.php?id=1>>.

Com o passar dos anos, a quantidade de informação disponível para realização de projetos e estudos tem sido cada vez mais substancial, principalmente no campo meteorológico, onde as séries históricas das variáveis são cada vez maiores. Pelo exposto, é interessante desenvolver técnicas de redução, classificação e filtragem destas informações. Uma técnica estatística multivariada muito utilizada é a análise de componente principal (ACP), essa técnica tornou-se popular na análise de dados atmosféricos após o artigo de Lorenz (1956), que denominou as funções de transformação de base como Funções Ortogonais Empíricas (FOE). Ambos os nomes são comumente utilizadas, e referem-se ao mesmo conjunto de procedimentos (WILKS, 2006).

Preisendorfer (1988) utilizou essa técnica porque é um método de redução de dados, tendo como vantagem o fato de que a máxima variância pode ser explicada pela classificação

dos autovetores associados aos maiores autovalores da matriz de correlação, ou seja, a amostra de dados originais pode ser analisada a partir de um número pequeno de componentes independentes entre si (ortogonais).

A ACP consiste essencialmente em reescrever as coordenadas das amostras em outro sistema de eixo mais conveniente para a análise. Em outras palavras, as  $n$  variáveis originais geram, através de suas combinações lineares,  $n$  componentes principais, cuja principal característica, além da ortogonalidade, é que são obtidos em ordem decrescente de máxima variância, ou seja, a primeira componente principal detém mais informação estatística que a segunda componente principal, que por sua vez tem mais informação estatística que a terceira componente principal e assim por diante. Este método permite a redução da dimensionalidade dos pontos representativos das amostras (Moita Neto e Moita, 1998).

A análise de agrupamentos dos dados também foi realizada. Esta técnica estatística de análise multivariada tem como objetivo principal classificar os indivíduos de uma população que são conhecidos por suas características, em grupos que sejam homogêneos intragrupos e heterogêneos intergrupos. As técnicas da Análise de Agrupamento podem ser hierárquicas e não-hierárquicas. Na primeira são produzidas sequências de partições em classes cada vez mais vastas, e nas não-hierárquicas efetua-se uma partição em um número de classes fixado a “priori”. Aqui foi utilizado o método hierárquico de Ward (WILKS, 2006)

Sansigolo e Nery (1998) aplicaram a análise dos fatores principais a 31 anos de dados mensais (1960-1991), de 47 estações meteorológicas distribuídas sobre a região sul e sudeste do Brasil, obtendo sua regionalização e melhor compreensão de seus padrões sazonais. A região apresentou grande variabilidade sazonal, com ciclo anual bem definido, com o primeiro fator comum temporal explicando 52% e o segundo 33% da variância total. Nesta perspectiva foram identificados 3 grupos homogêneos em relação às precipitações anuais e distribuição sazonal, sendo um grupo na região sudeste e dois grupos mais dispersos na região Sul.

Romero et al. (1998b) usaram uma base de dados homogeneizados de precipitação diária em 410 locais por 30 anos (1964-1993), para produzir mapas anuais e sazonais, caracterizando, os montantes acumulados, frequências, intensidade e persistência ou ausência de chuvas na região. Essa análise foi aplicada separadamente para três décadas consecutivas, 1964-1973, 1974-1983 e 1984-1993.

Assim, o principal objetivo deste trabalho consiste em utilizar a técnica da ACP para reduzir os dados, definindo os padrões de variabilidade espacial e temporal das precipitações

na região do Estado do Piauí. E identificar comportamentos homogêneos de determinadas regiões do Piauí através do método de agrupamento hierárquico aglomerativo de Ward.

## 2. MATERIAL E MÉTODOS

### 2.1 MATERIAL

Os dados foram fornecidos pela Agência Nacional de Águas (ANA), em forma de valores mensais de precipitação para 27 estações pluviométricas do estado do Piauí, englobando a série histórica de 1963 a 2000. Observe na Tab.1 a relação das estações com suas respectivas localizações no mapa do Estado do Piauí. Estes dados foram tratados e organizados para se calcular as médias mensais e anuais para cada estação. Esta planilha foi utilizada no programa SPSS 13.0 (*Statistical Package for the Social Sciences*), para fazer a análise fatorial em componentes principais e o método de agrupamento hierárquico de Ward. Utilizou-se também o programa Surfer 7.0 na produção dos mapas.

**Tabela 1.** Espacialização das estações (coluna a) e estações selecionadas e siglas (coluna b)

Distribuição espacial das estações		
		1 Lu. Co Luiz Correia
		2 Porto Porto
		3 Pirac Piracuruca
		4 Barras Barras
		5 Pedro Pedro II
		6 Jô. Fr. José de Freitas
		7 Teres Teresina-Chesf
		8 Alto L. Alto Longo
		9 Cas. Pi Castelo do Piauí
		10 Amar Amarante
		11 Val. Pi Valência do Piauí
		12 Oeir Oeiras
		13 Pic Picos
		14 La. Sa <u>Landri Sales</u>
		15 S.F.Pi São Francisco do Piauí
		16 Jaic Jaicos
		17 Flores Flores do Piauí
		18 Si. Me Simplicio Mendes
		19 Paulis Paulista
		20 B.do Pr <u>Barra do Prata</u>
		21 S. J. Pi São Joao do Piauí
		22 Cri. Ca Cristino Castro
		23 S.R.No São Raimundo Nonato
		24 Ac. Car Acude Carocal
		25 Mo.Ale Monte Alegre
		26 AveLop Avelino Lopes
		27 Corr Corrente

## 2.2 - MÉTODOS

### 2.2.1 - Análise Fatorial em Componentes Principais (ACP)

A partir da matriz de dados  $X$  ( $n$  indivíduos  $\times$   $p$  caracteres), cujas variáveis possuem média  $E(x_j)=m(j)$  e variância  $V(x)=s(j)$ ,  $j= 1...p$ , pode-se obter a matriz de variância e covariância  $S$ , sendo assim:

$$S = \frac{1}{n} X \cdot X^t \quad (1.1)$$

Considerando a matriz  $X$  com valores centrados e  $X^t$  sua matriz transposta e  $n$  é o número de indivíduos ou estações. Normalizando a matriz das variáveis  $x_1, x_2, x_3, \dots, x_p$  normalizadas (média zero e variância um) através da média e do desvio padrão, a matriz variância e covariância será igual à matriz de correlação  $R$ , dada por:

$$R = \frac{1}{n} Z \cdot Z^t \quad (1.2)$$

Sendo a matriz  $R$  uma matriz de correlação simétrica e positiva de dimensão ( $p \times p$ ), será diagonalizável por uma matriz ortogonal  $A$ , de mudança de base denominada de matriz dos autovetores. Com isso a matriz diagonal  $D$ , cujos os elementos diagonais são os autovalores de  $R$ , é obtida por:

$$D = A^{-1} \cdot R \cdot A = A^t \cdot R \cdot A \quad (1.3)$$

onde  $A^{-1}$  é a inversa da matriz  $A$ .

As CP's  $U_1, U_2, \dots, U_p$ , são obtidas por combinações lineares entre a transposta dos autovetores de  $A^t$  e a matriz de observações  $X$ , isto é:

$$U = A^t \cdot X \quad (1.4)$$

Os valores  $X_i$  do  $n$ -ésimo local podem ser estimados pela seguinte expressão:

$$X_i = A_{j1}U_1 + A_{j2}U_2 + \dots + A_{jk}U_k + \dots + A_{jp}U_p \quad (1.5)$$

onde as variáveis de  $U_p$  é o conjunto dos autovetores de  $X$  ordenados por ordem decrescente dos autovalores mais significativos de  $a$  em  $a_k$ , qual seja:

$$U_k = \sum_{j=1}^n a_{jk} X_k \quad (1.6)$$

O melhor ajuste das componentes é feito através da explicação da variância, pela ordem crescente dos autovalores. A soma dos valores próprios é a soma das variâncias centradas e normalizadas. Assim, a porcentagem da variância dos autovalores em ordem decrescente que é dada por:

$$\%Var(X) = \frac{\sum_{i=1}^q \lambda_i}{\sum_{j=1}^p \lambda_j} \quad (1,7)$$

O método de rotação ortogonal usado foi o VARIMAX, que maximiza a variância em cada componente. A correlação entre a i-ésima variável original e a i-ésima componente principal é dada por:

$$Coor(X_i, Z_j) = a_{ij} \cdot \sqrt{\lambda_i} \quad (1,8)$$

$a_{ij}$  é o j-ésimo elemento do i-ésimo autovetor e  $\lambda_i$  o i-ésimo autovalor.

### 2.2.2 – Métodos de Agrupamentos (Cluster)

Existem dois tipos de métodos ou algoritmos de classificação de grupos. Um é o método hierárquico, onde a partição dos grupos se dá a partir de um mínimo de grupos não definidos inicialmente, os grupos majoritários são divididos em sub-grupos minoritários agrupando aqueles indivíduos que apresentam características semelhantes. A estrutura final das classes é apresentada sobre a forma de uma árvore de classificação (dendrograma) que apresenta uma síntese objetiva dos resultados. O outro é o método não hierárquico de classificação onde o número dos grupos é definido a priori. Nos dois métodos de agrupamentos a classificação dos indivíduos em grupos distintos é feita a partir de uma função de agrupamento e de um critério matemático de agrupamento (EVERITT, 1993; WILKS, 2006).

### 2.2.3 - Método de Ward

Este é um método hierárquico que utiliza a distância euclidiana para medir a similaridade ou dissimilaridade entre os indivíduos, ou seja, a distância entre os indivíduos  $X_i$  e  $X_j$  é dada por (EVERITT, 1993):

$$d_{ij} = [X_i - X_j] = \left[ \sum_{k=1}^p (X_{i,k} - X_{j,k})^2 \right]^{1/2} \quad (1.9)$$

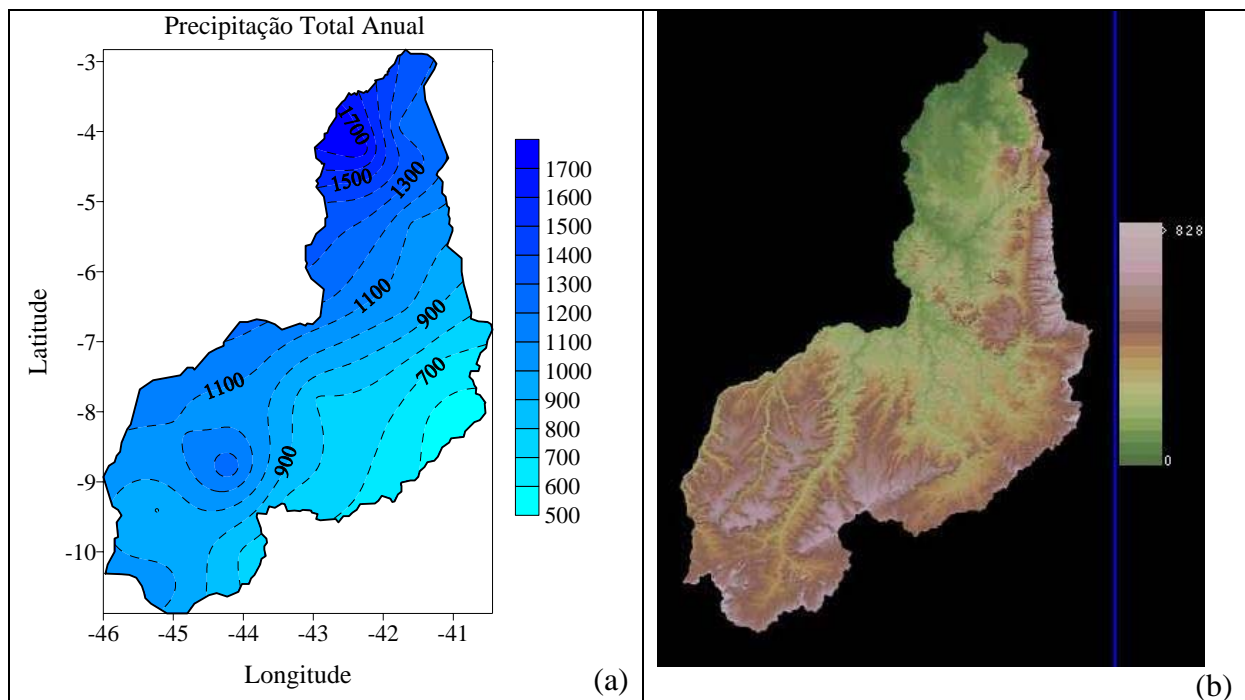
O método de Ward propõe que em qualquer fase da análise a perda de informação que resulta do agrupamento de elementos entre grupos, é medida pela Soma dos Quadrados dos Desvios (SQD) de cada ponto a média do grupo a qual pertence. (EVERITT, 1993).

$$SQD = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \quad (2.0)$$

onde  $n$  é o número total de elementos do agrupamento e  $x_i$  é o  $i$ -ésimo elemento do agrupamento.

### 3. RESULTADOS E DISCURSSÃO

O programa SPSS encontrou dois fatores mais significativos que explicam 88% da variância dos dados, como podemos ver na Tab. 2. Para facilitar a explicação, foram plotados no programa Surfer os gráficos de precipitação acumulada e dos dois fatores comuns. A Fig.1 mostra a variabilidade espacial dos totais médios anuais da precipitação no Estado do Piauí.



**Figura 1.** Distribuição da precipitação média anual (a), relevo estado do Piauí (b).

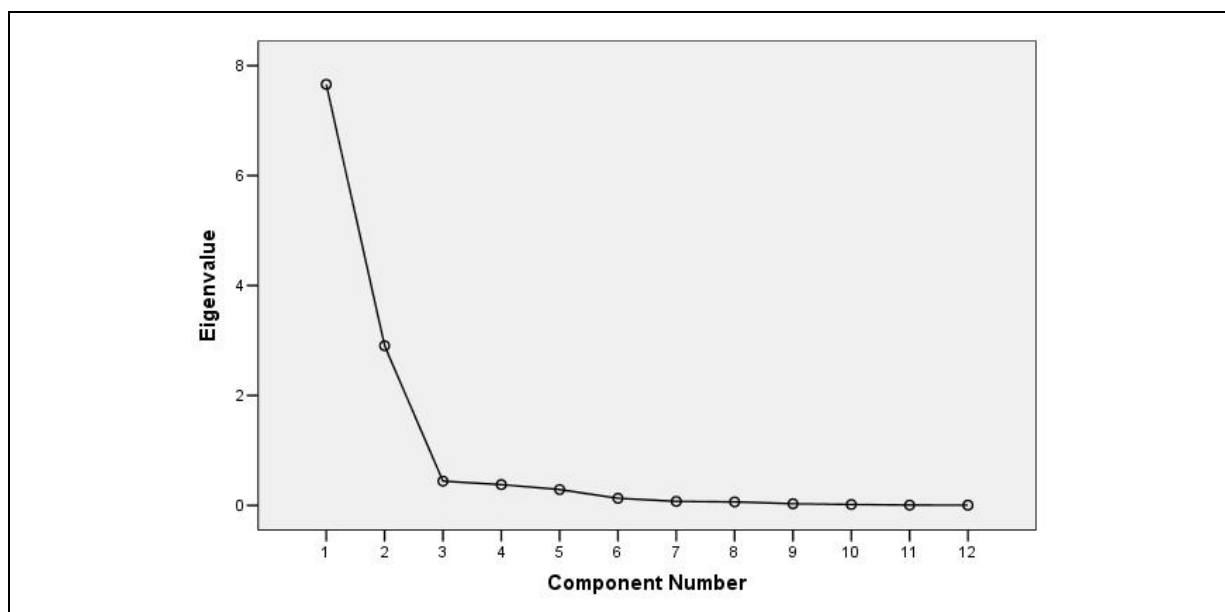
Analisando a Figura 1a, os maiores valores de precipitação anual se concentram na parte norte do estado, acima de  $5,5^{\circ}\text{S}$ , com núcleo máximo no lado leste sobre o Rio Parnaíba de 1700mm. Os menores valores estão na parte central do estado, região semi-árida. A topografia vista na figura 1b, ajuda bastante na explicação dessa configuração, pois a região norte, parte mais baixa, favorece a atuação de sistemas meteorológicos como veremos adiante. E a região central, cercada de serras, principalmente do lado leste, dificultado o escoamento dos sistemas de leste e a convergência de umidade, aumentando a sensação térmica e deixando o tempo abafado. Na parte sul do estado, destaca-se o comportamento da isoeta de 1000 mm, que segue o vale contornando a região montanhosa, comprovando a influencia da topografia.

A Tab. 2 mostra o peso de cada variável, representada pelos meses do ano, em relação ao conjunto de dados médios mensais da precipitação pluvial. Foram considerados os dois primeiros fatores comuns rotacionados, tanto pelo critério *scree plot* como pelo critério de truncamento de Kaiser que serão explicados a seguir.

**Tabela 2.** Total da variância explicada

Componente	Autovalores			Cargas			Cargas Rotacionadas		
	Total	Variância (%)	Acumulada (%)	Total	Variância (%)	Acumulada (%)	Total	Variância (%)	Acumulada (%)
1	7,659	63,824	63,824	7,659	63,824	63,824	6,965	58,045	58,045
2	2,904	24,202	88,026	2,904	24,202	88,026	3,598	29,981	88,026
3	0,442	3,681	91,707						
4	0,379	3,155	94,862						
5	0,289	2,411	97,274						
6	0,130	1,085	98,358						
7	0,074	0,614	98,973						
8	0,064	0,535	99,508						
9	0,031	0,259	99,766						
10	0,017	0,142	99,908						
11	0,006	0,051	99,959						
12	0,005	0,041	100,000						

O gráfico abaixo é a curva *scree plot* que relaciona o número de ordem de cada eixo com o valor próprio que lhe está associado (Fig. 2). Se essa curva evidenciar uma estabilização dos valores próprios, pode-se reter apenas os eixos com números de ordem superiores aquele que inicia a estabilização. Traçar o espectro de autovalores com uma escala linear vertical produz o que é conhecido como o gráfico *scree*.



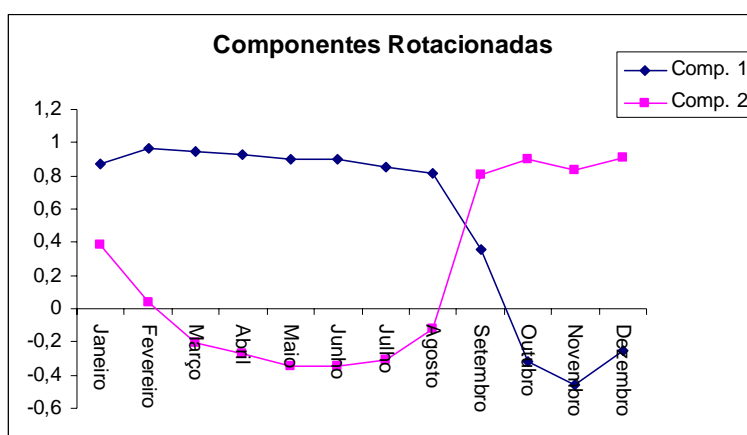
**Figura 2.** Scree Plot



Ao utilizar o gráfico *scree* qualitativamente, o objetivo é localizar um ponto que separa uma parte de forte inclinação para a esquerda, de outra parte com inclinação mais rasa para a direita. O número de componentes principais em que a separação ocorre é então considerado como o corte de truncamento, o critério inclinação *scree* não envolve inferência estatística quantitativa (WILKS, 2006).

Assim, o número escolhido de Componentes Principais foi baseado no critério de truncamento de Kaiser, que considera como mais significativos os autovalores cujos valores sejam superiores à unidade (GARAYALDE et al., 1986).

Na Fig. 3, analisando a componente 1 se observa sua melhor correlação com os meses de janeiro a agosto possuindo valores acima de 0.8. Enquanto que a componente 2, se correlaciona bem com valores acima de 0.8 para os meses de setembro a dezembro. Há uma pequena correlação de 0.4 da componente 2 para o mês de janeiro e da componente 1 para o mês de setembro. Em outras palavras a primeira componente explica os eventos de precipitação ocorridos desde meados do verão até o final do inverno e por isso ele justifica 63,8% do dados não rotacionados e 58% com rotação. Enquanto que a segunda componente explica os eventos da primavera e início do verão explicando 24,2% dos dados sem rotação e 30% com rotação. Como pode ser visto na tabela 2, as duas componentes juntas, ultrapassam 88% da variância dos dados.



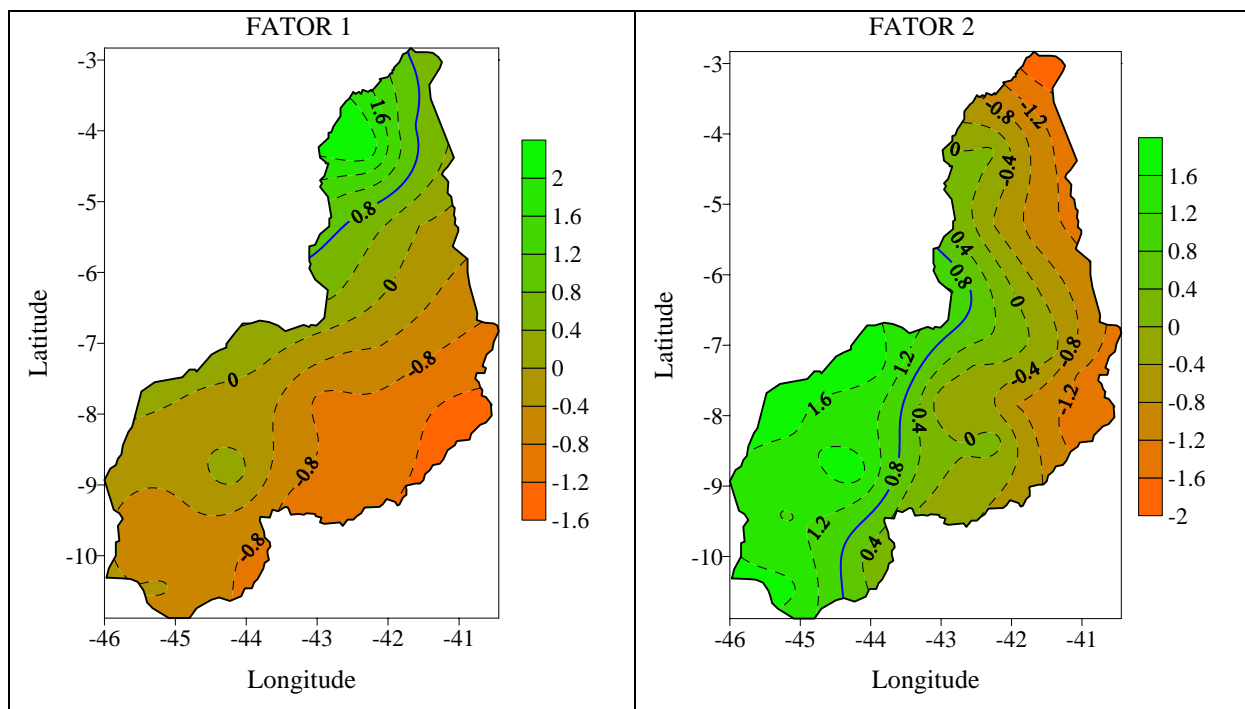
**Figura 3.** Distribuição temporal das componentes principais.

Durante o desenvolvimento do trabalho, fez uma análise com três componentes, o comportamento das duas primeiras foi similar ao mostrado acima, a terceira componente mostrou um pico no mês de agosto, que é o mês mais seco de todo o Estado, mais sem significância para todos os outros meses. O gráfico do Fator 3, não teve representação espacial, sem peso significativo pra nenhuma região e justamente pela falta de

representatividade espacial e temporal e também pelo critério de Kaiser adotado, neste trabalho nos retemos as duas primeiras componentes.

A diferença básica neste trabalho entre os termos fatores e componentes é que os fatores explicam a distribuição espacial da variável formando uma nova base hierarquizada do espaço em estudo. Enquanto que as componentes explicam a distribuição temporal da variável.

No mapa do fator 1 (Fig. 4) se observa que os maiores pesos se encontram na parte norte do estado, justamente na mesma área das maiores precipitações, e os menores pesos na região de menor precipitação. É interessante notar que as isolinhas do fator 1 praticamente sobrepõe as isoetas no gráfico da precipitação (Fig. 2) Os valores acima de 0,8 foram considerados com pesos mais significativos para se determinar a região de melhor explicação do fator. Assim o fator 1 se associa bem com a distribuição de chuva do norte do estado (cor verde), e não explica a precipitação ocorrida na parte sudeste (cor laranja). Para o fator 1, os sistemas meteorológicos mais significativos são a ZCIT (Zona de Convergência Intertropical) e as brisas marítimas que conseguem penetrar bastante no continente do lado oeste devido a planície do Rio Parnaíba.

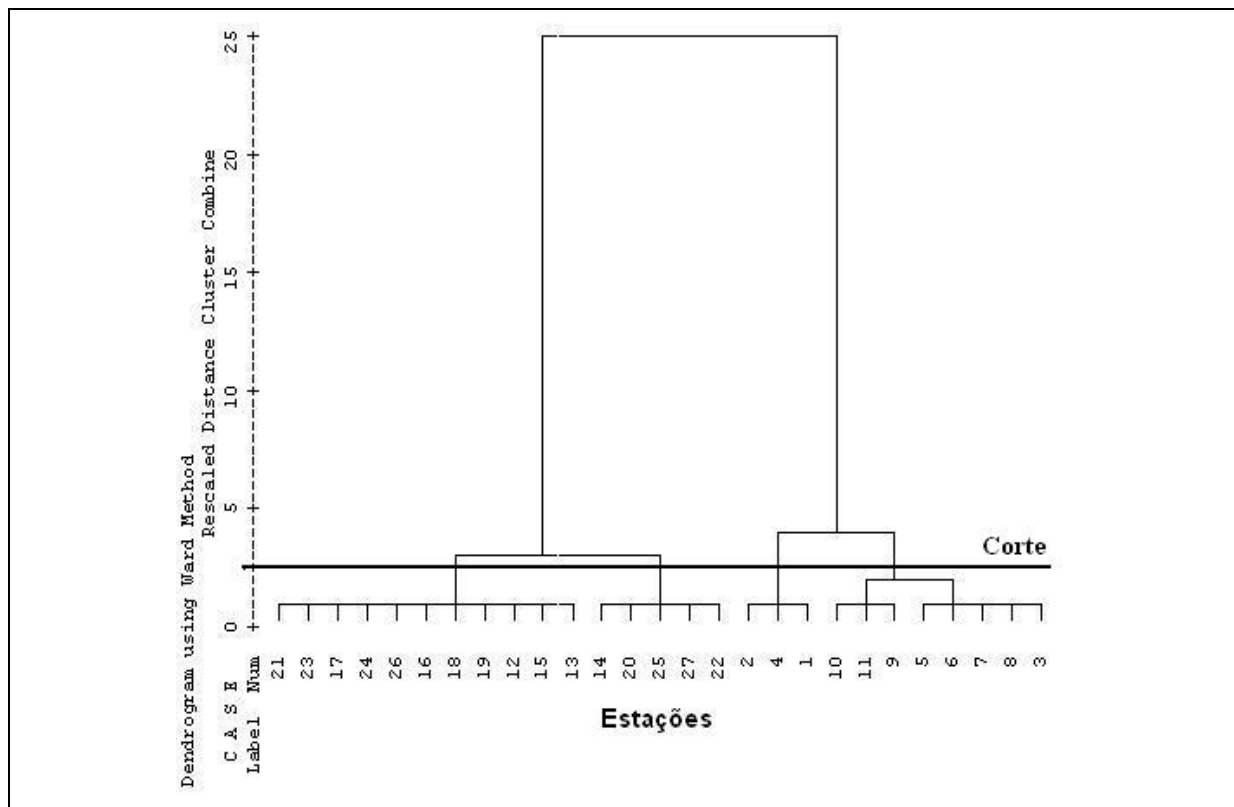


**Figura 4.** Distribuição espacial dos fatores 1 e 2.

Analogamente, o fator 2 (Fig. 4) se relaciona bem com a parte sudoeste do estado (cor verde), onde a precipitação média anual varia entre 900mm e 1200mm. O fator 2 não conseguiu explicar a precipitação do leste do Piauí (cor laranja). Já para o fator 2, o sistema que melhor justifica esse comportamento é a convergência de umidade oriunda da floresta

amazônica e do Atlântico Norte, que é responsável pelo clima úmido da região Norte. Um outro sistema responsável pela umidade no sudeste do Piauí é a ZCOU (Zona de Convergência de Umidade no sentido sudeste/noroeste), esse escoamento para sudeste, as parcelas se elevam conforme o conceito de esteira transportadora úmida, gerando uma extensa banda de nebulosidade no Brasil Central, se essa convergência recebe energia e massa de sistemas frontais estacionários sobre o Oceano Atlântico Sul, pode se manter em uma escala temporal maior, recebendo o nome de ZCAS, que aumenta ainda mais a umidade.

Um outro sistema que interfere na precipitação em todo o Estado do Piauí, ou seja, em ambos os fatores, é o VCAN (Vórtice Ciclônico dos Altos Níveis), com atuação durante o verão do H.S. Além dos sistemas locais, gerados pelos reservatórios d'água e pelas montanhas, provocando as brisas lacustres e as brisas vale-montanha respectivamente.

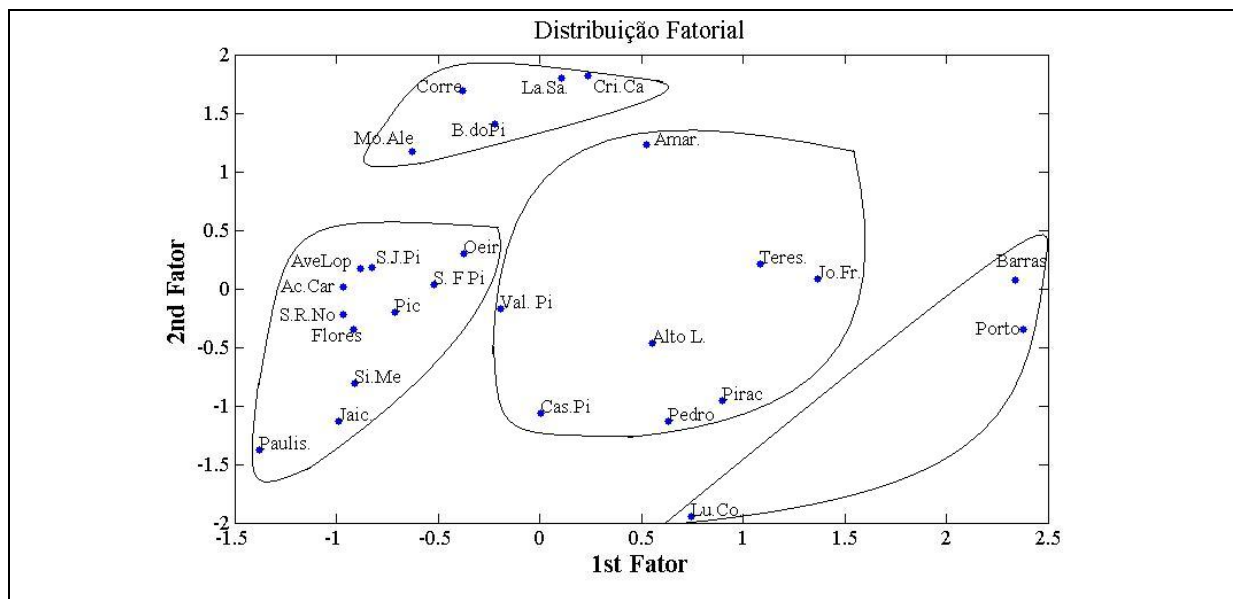


**Figura 5:** Dendrograma resultante do método de Ward para as 27 estações pluviométricas utilizadas.

A seguir, damos entrada na análise de agrupamento pelo método hierárquico de Ward, a Fig. 5 mostra o dendrograma representativo da homogeneidade existente entre as precipitações de cada estação. Um corte na horizontal, delimitando a formação de 4 grupos intra-homogêneos se mostrou mais eficaz e melhor distribuído que um outro corte. Um corte com dois grupos dividiria o estado em parte norte e sul, delimitando as configurações

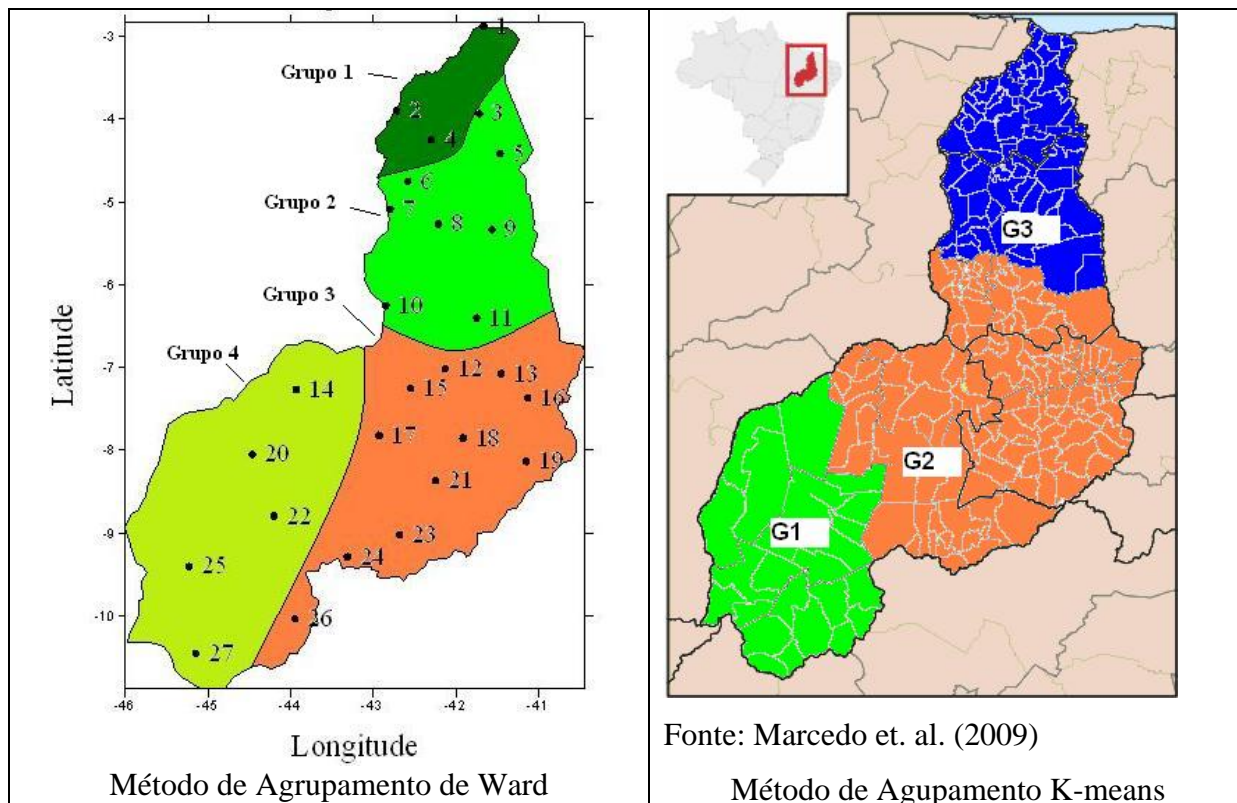
específicas da distribuição de chuva. De forma que a melhor caracterização para método de Ward foi com 4 grupos.

A Fig. 6, nos fornece o gráfico do fator 1 versus fator 2, onde é possível detectar o posicionamento das estações em relação ao peso dos dois fatores, podemos ver também os grupos detectados pelo método de Ward, mostrando que as estações de cada grupo possui pesos próximos para os dois fatores. Assim, este gráfico fornece uma janela privilegiada (estatisticamente) para observação dos pontos no espaço n-dimensional. Para identificação e localização das estações consulte a tabela 1 (coluna b).

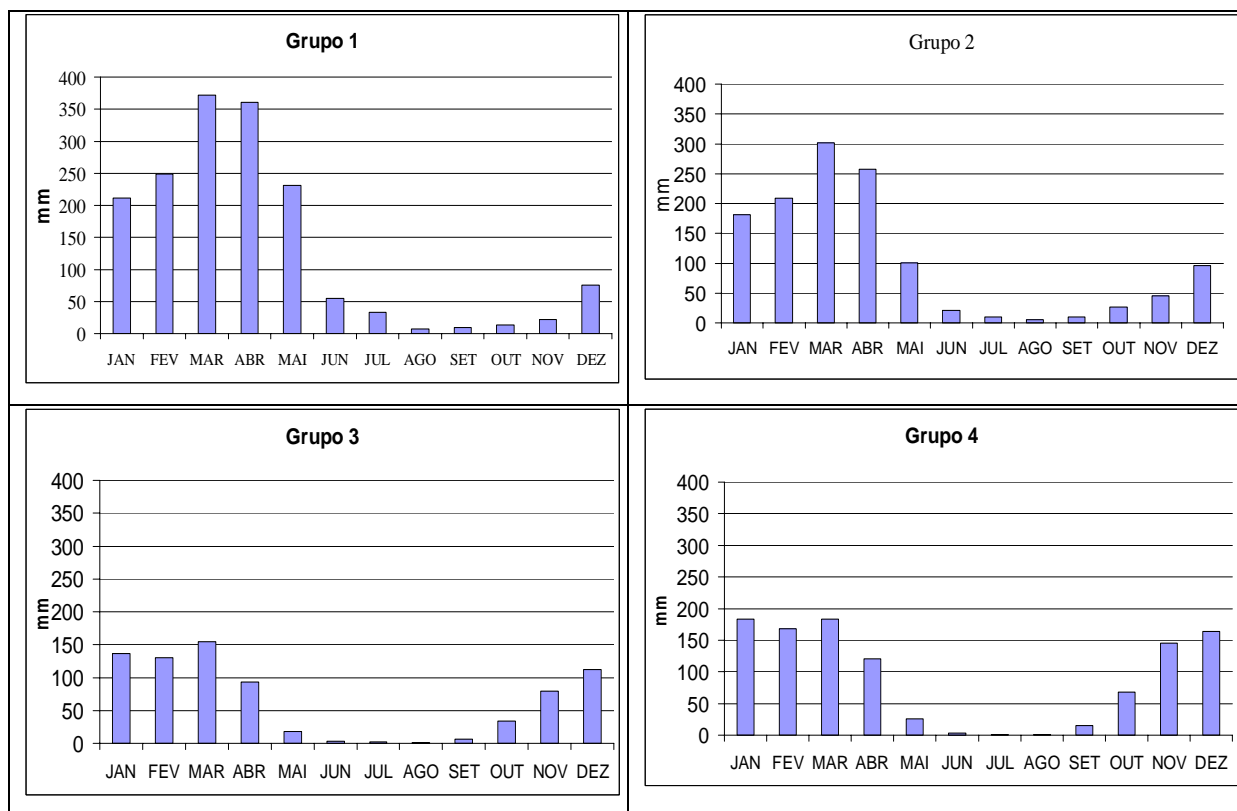


**Figura 6.** Distribuição dos 4 grupos em relação aos dois fatores.

A figura 7 apresenta a distribuição geográfica dos grupos no Estado do Piauí, onde o grupo 1 (verde escuro) se localiza mais a norte, em seguida vem o grupo 2 (verde claro). Na parte central está o grupo 3 (laranja) e no sudoeste/sul se encontra o grupo 4 (amarelo escuro). Um artigo que utilizou um método não hierárquico para o Piauí, apresentou uma boa configuração com três grupos, utilizando o método K-means (Marcedo et. al. 2009). Embora deva existir diferenças nas série histórica de precipitação e nas estações utilizadas pelos dois artigos, se verifica que existe uma similaridade entre os grupos apresentados pelos dois métodos, a principal diferença é que o método de Ward manteve um grupo a mais no norte (grupo 1), isso devido os altos valores de precipitação média em relação a média do outro grupo 2. Uma outra diferença seria pequenos desvios nas posições geográficas dos grupos, mais é importante resaltar que o método de Ward define sozinho a quantidade de grupos, podendo modificar apenas o corte no dendrograma. Enquanto que no método K-means a quantidade de grupos é escolhida previamente.



**Figura 7.** Distribuição espacial dos grupos de precipitação segundo o método de Ward.



**Figura 8.** Precipitação média mensal em cada grupo.

Observando a figura 8, verificamos que a configuração mensal da precipitação é similar entre os grupos 1 e 2, onde os meses mais chuvosos são de janeiro a maio, a diferença entre os grupos está apenas na intensidade da precipitação, no grupo 1 os valores médios são maiores em todos os meses. A mesma análise pode ser feita para a configuração similar entre os grupos 3 e 4, onde os meses mais chuvosos vão de novembro a abril, porém com valores bem menores que nos dois primeiros grupos. Na relação entre os grupos 3 e 4, se verifica que no grupo 4 os valores médios mensais são maiores que os do grupo 3. Assim, podemos entender que os sistemas meteorológicos referidos anteriormente ao fator 1, atual nos grupos 1 e 2, enquanto os fenômenos relacionados ao fator 2, atua nos grupos 3 e 4.

#### **4. CONCLUSÃO E SUGESTÃO**

A estação chuvosa do Estado vai de dezembro a abril com valores médios acima de 112mm mensais e menores que 230mm. Detectou-se diferença entre os regimes de chuva do norte e do sul, tanto na ocorrência, quanto na intensidade. Indicando assim, atuação de sistemas meteorológicos diferentes. O máximo de precipitação ocorre em março, e o mínimo em agosto, tanto nos regimes do norte como do sul do Piauí. No sul do Estado, o chamado grupo 4, foi bem representado pelo Fator 2, enquanto que o norte do Estado foi representado pelo Fator 1. Ambos os fatores obtiveram pesos negativos na parte mais seca do semi-árido, dentro do grupo 3. O grupo 2, ficou como área de transição dos dois fatores.

Caso haja interesse em estudar melhor os grupos 2 e 3, talvez seja interessante fazer uma análise com 4 fatores. No desenvolvimento do trabalho fizemos com três fatores, mais o terceiro se mostrou inconclusivo, possuindo pesos baixos para quase todo o Estado, a terceira componente mostrou uma inclinação acentuada no mês de agosto, que é o mês mais seco em todos os grupos. Mais isto não foi exposto nos resultados pelo critério adotado no trabalho, sendo a componente com autovalor menor que 1 e variância de 3%. O método de Ward dividiu bem a região em 4 grupos fechados e homogêneos facilitando a explicação e entendimento. Uma possível divisão em 3 grupos não ficou bem configurado, possuindo integrantes de grupos dispersados.

#### **5. REFERÊNCIAS BIBLIOGRÁFICAS**

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Sistema de informações hidrológicas (HidroWeb). Disponível em: < <http://hidroweb.ana.gov.br/>>. Acesso em 26 abr. 2010.

Aspectos Naturais do Piauí. Disponível em: <<http://www.brasilecola.com/brasil/aspectos-naturais-piaui.htm>>. Acesso em 26 abr. 2010.

BUSSAB, W. O.; MIAZAKI, E. S.; ANDRADE, D. **Introdução à análise de agrupamentos**. São Paulo, Ass. Bras. de Estatística, 1990. 105p.

Características físicas do Piauí. Disponível em: <<http://www.pi.gov.br/piaui.php?id=1>>. Acesso em 26 abr. 2010.

DANIEL S. WILKS. **Statistical Methods in the Atmospheric Sciences**, 2nd Ed. v. 91 p. 463-507, 2006.

EVERITT, B.S. **Cluster analysis**. Heinemann Educational Books, London: Academic Press, 3ª edição, 1993. p.170.

GARAYALDE, E.J.G.; da SILVA, M.G.A.; TAVARES, A. de SÁ. Classificação meso-climática da região Sul do Brasil pela análise de componentes principais. in: CONGRESSO INTERAMERICANO DE METEOROLOGIA., Brasília 1986. Anais CONGRESSO BRASILEIRO DE METEOROLOGIA, Brasília: Sociedade Brasileira de Meteorologia. v. 1, p. 119-124. 1986.

LORENZ, E.N. Empirical Orthogonal Functions and Statistical Weather Prediction. Sci. Rep. nº 1. Statistical Forecasting Project, Department of Meteorology, Cambridge, Massachusetts (EUA), MIT, 49 p. 1956.

MARCEDO, M. J. H.; BRAGA, C. C.; DANTAS, F. R. C.; SILVA, E. D. V. **Variabilidade espacial e da pluviometria no Estado do Piauí**. 3º Simp. Inter. de Climatologia. Canela: Sociedade Brasileira de Meteorologia, outubro, 2009.

MOITA NETO, J. M.; MOITA, G. C. Uma introdução à análise exploratória de dados multivariados. **Química Nova**, v. 21, p. 467- 469, 1998.

NEVES, D. G. Variabilidade interanual do índice de vegetação no leste do Nordeste do Brasil no período de 1982-1999. **Dissertação** (Mestrado em Meteorologia) – Universidade Federal de Campina Grande, 2005.

PREISENDORFER, R. Principal component analysis in meteorology and oceanography. Amsterdam: **Elsevier**, 425 p. 1988.

ROMERO, R., GUIJARRO, J.A., RAMIS, C. AND ALONSO, S. 1998b. ‘A 30 year (1964–1993) daily rainfall data base for the Spanish Mediterranean regions: First exploratory study’, *Int. J. Climatol.*, 18, 541–560.

SANSIGOLO, C.; NERY, J. T. **Análise de fatores comuns e agrupamentos das precipitações nas regiões Sudeste e Sul do Brasil**. In: X CONGRESSO BRASILEIRO. DE METEOROLOGIA, Brasília, 1998.

WARD, J. H. **Hierarchical grouping of optimize an objective function**. J. American Statistical Association, 58:236-244, 1963.