

# Análise de componentes principais com dados pluviométricos no estado do Ceará

Maytê Duarte Leal Coutinho<sup>1</sup>, Michelyne Duarte Leal Coutinho<sup>2</sup>, José Ivaldo B. de Brito<sup>3</sup>

<sup>1</sup>Aluno de mestrado da Universidade Federal de Campina Grande – UACA/CTRN/UFCG  
[mayteleal@hotmail.com](mailto:mayteleal@hotmail.com)

<sup>2</sup>Aluno de doutorado do Instituto Nacional de Pesquisas Espaciais – CPTEC/INPE  
[michelyne.coutinho@cptec.inpe.br](mailto:michelyne.coutinho@cptec.inpe.br)

<sup>3</sup> Professor da Universidade Federal de Campina Grande – UACA/CTRN/UFCG  
[ivaldo@dca.ufcg.edu.br](mailto:ivaldo@dca.ufcg.edu.br)

## **ABSTRACT:**

The objective this work is to verify if the precipitation on the Ceará, that presents a strong variability in your space distribution, it can be represented for a reduced number of components principal. Therefore, the technique of principal components analysis (PCA) was used applied to the monthly data of thirty meteorological stations in the period from 1970 to 2000, distributed on the State. It was verified that is just necessary the first three components to explain the variability of the space distribution of the rains on the Ceará.

**Key Words:** PCA, precipitation, climate

**Palavras-Chave:** ACP, precipitação, clima

## **1. INTRODUÇÃO**

O estado do Ceará situa-se no norte do NEB e apresenta clima predominantemente semi-árido. Em pelo menos oito meses do ano, chove pouco, dependendo da localidade, as precipitações podem variar de menos de 500 mm/ano, em algumas áreas do Sertão do Inhamuns, a mais de 1.400/ano, no litoral de Fortaleza, a valores em torno de 1.000 mm/ano, na Chapada do Araripe, e a mais de 1.800 mm/ano em áreas da Serra da Ibiapaba. Devido à grande variabilidade espacial da precipitação são vários os sistemas meteorológicos atuantes na produção de chuvas no Estado, os principais são: Zona de Convergência Intertropical (ZCIT), Ondas de Leste, Vórtice Ciclônicos de Altos Níveis (VCAN) e sistemas de mesoescala, como linhas de instabilidade e complexos convectivos de mesoescala.

Por outro lado, o estudo da climatologia de uma área específica sem ter um único ponto de observação é feito, em geral, por estimativa dos elementos climáticos de áreas adjacentes. As estimativas, capazes de representar o clima da área sem informação, podem ser feitas por técnicas estatísticas. E mais utilizada em diversas áreas do conhecimento é análise multivariada e uma das técnicas mais utilizadas dentro da análise multivariada é análise de componentes Principais (ACP).

Portanto, o presente trabalho tem como exercício básico verificar se para regiões com variáveis climáticas apresentando uma grande variabilidade espacial, como é o caso da distribuição da precipitação sobre o Ceará, podem ser analisada através de poucas componentes principais (3 a 5), ou se são necessário uma gama muito elevada de componentes (mais de 10). Ressalta-se que a análise climática representada por um grande número de componentes pode levar a resultados imprecisos, uma vez que as componentes de ordem maior podem ser decorrentes da combinação linear das primeiras componentes (Johnson & Wichern, 1998).

## 2. DADOS E METODOLOGIA

Neste estudo são usados dados pluviométricos de trinta estações do estado do Ceará, adquiridos da Agência Nacional das águas (ANA) para o período de 1970 a 2000. Foram obtidas as precipitações mensais e anuais com estes dados e a técnica de análise das componentes principais (ACP) foi usada para definir as regiões homogêneas de precipitação pluvial no estado do Ceará. Utilizou-se o software SPSS em sua versão 9.0 para processamento dos dados pluviométricos mensais.

### 2.1- Análise de componentes principais

A Análise de Componentes Principais (ACP) foi desenvolvida por PEARSON (1901), é uma técnica estatística de análise multivariada que busca, através de transformações lineares, descrever a interrelação das variáveis em estudo, explicando estrutura de interdependência dessas e criando, assim, um novo conjunto com o número de variáveis inferior ao conjunto original. Para o cálculo das Componentes Principais (CP) de um conjunto de dados, inicialmente deve-se obter os autovalores e os autovetores a partir da matriz de correlação ou da matriz de variância-covariância entre as variáveis desse conjunto. A aplicação da ACP a um conjunto de dados de grande dimensão é interessante, inicialmente, apenas para determinar combinações lineares das variáveis originais que expliquem o máximo possível a variação existente nos dados iniciais (Silva, 2001). Por exemplo, para  $m$  componentes e  $p$  variáveis ( $p \geq m$ ), tem-se a seguinte configuração de variáveis latentes:

$$\begin{aligned} CP_1 &= a_{11} X_1 + a_{21} X_2 + \dots + a_{p1} X_p \\ CP_2 &= a_{12} X_1 + a_{22} X_2 + \dots + a_{p2} X_p \\ &\vdots \\ CP_m &= a_{1m} X_1 + a_{2m} X_2 + \dots + a_{pm} X_p \end{aligned}$$

A solução deste sistema de equações considera a variação total presente no conjunto das variáveis iniciais, onde a primeira  $CP_1$  explica o máximo possível da variância dos dados iniciais, enquanto a segunda  $CP_2$  explica o máximo possível da variância ainda não explicada e assim por diante, até a última  $CP_m$  que contribui com a menor parcela de explicação da variância total dos dados iniciais (Silva, 2001).

A variância total do sistema ( $V$ ) é definida como a soma das variâncias das variáveis observadas, assim  $V$  é dada por:  $V = \text{traço } S = \sum_{i=1}^p S_{ii} = \sum_{i=1}^p \lambda_i$  em que  $S$  é a variância das variáveis observadas e  $\lambda_i$  são os autovalores.

A variância explicada por cada componente é:

$$\alpha_i = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \cdot 100 \quad (\%)$$

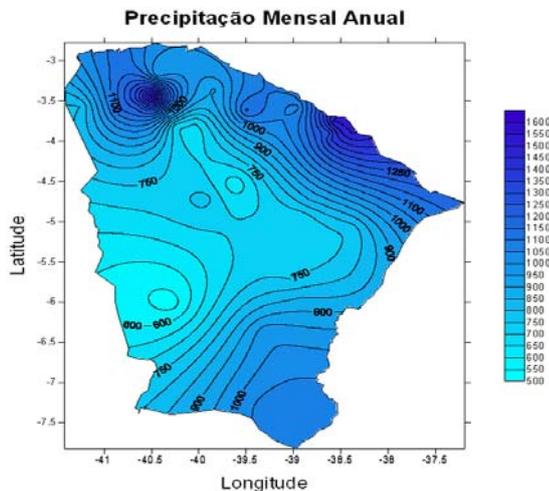
O número escolhido de Componentes Principais foi baseado no critério de Kaiser, que considera como mais significativos os autovalores cujos valores sejam superiores a um (Garayalde et al., 1986).

## 3. RESULTADOS E DISCUSSÕES

Os resultados são apresentados na forma de padrões da variabilidade temporal e espacial da precipitação no estado do Ceará, obtidos pelo método proposto na subseção 2.1.

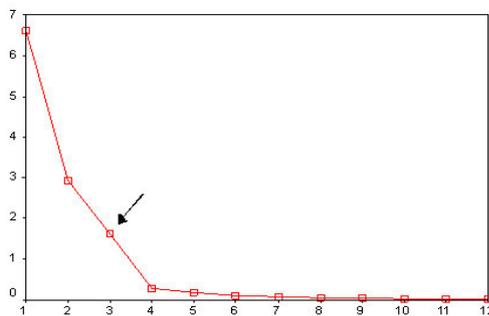
A Figura 1 mostra os totais médios anuais de precipitação no estado na região em estudo. No qual os maiores índices pluviométricos são vistos em grande parte da faixa

litorânea e leste da região, enquanto que os menores valores encontram-se na região do sertão (centro-oeste). Tal distribuição no norte do Ceará é caracterizada, principalmente, pela atuação da ZCIT no período de março a abril (Namias, 1972), já os totais de precipitação significativos na parte litorânea são resultados da atuação, das ondas de leste e de VCAN.



**Figura 3.** Distribuição espacial dos totais médios anuais da precipitação pluvial no estado do Ceará

Através do gráfico da curva de inércia, foi possível encontrar os fatores de significância das variáveis. A seta indica onde os valores foram truncados mostrando os três fatores de significância encontrados, pois aqueles menores que um (1,0) são considerados insignificantes, sendo assim, descartados.

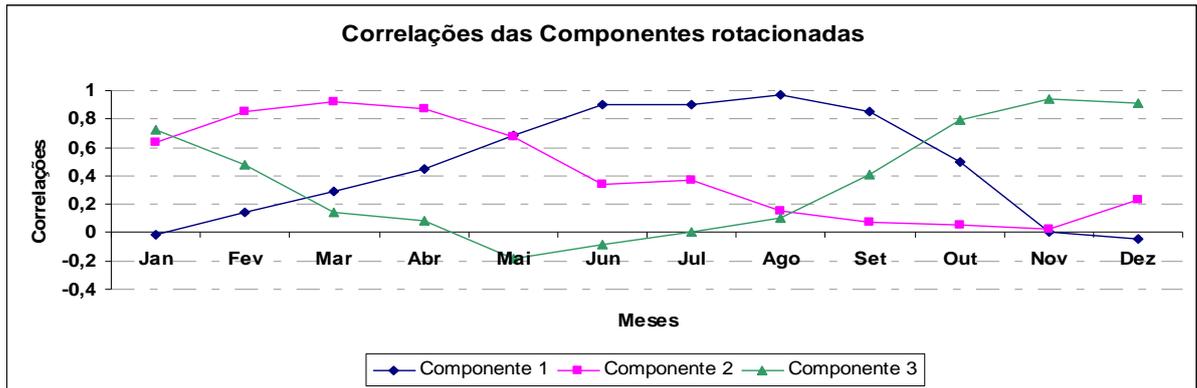


**Figura 2-** Curva de Inércia

A análise fatorial em componentes principais rotacionada, aplicada aos dados médios mensais da precipitação no Estado do Ceará produziram correlações sazonais e padrões espaciais associados a cada fator comum (pesos). A Tabela 1 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 três primeiros fatores comuns rotacionados, que explicaram 93,27% da variância total dos dados da precipitação pluvial, truncando os valores menores do que 1, segundo o critério de Kaiser.

**Tabela 1.** Seqüência dos autovalores na ordem decrescente e a contribuição das variâncias explicada (%) e total (%) dos dados médios mensais da precipitação pluvial rotacionados na região do Ceará.

| Fator     | Autovalor | Variância explicada | Variância acumulada |
|-----------|-----------|---------------------|---------------------|
| 1         | 6,62      | 55,20               | 55,20               |
| 2         | 2,94      | 24,51               | 79,72               |
| 3         | 1,63      | 13,56               | 93,28               |
| <b>12</b> | <b>0</b>  | <b>0</b>            | <b>100</b>          |

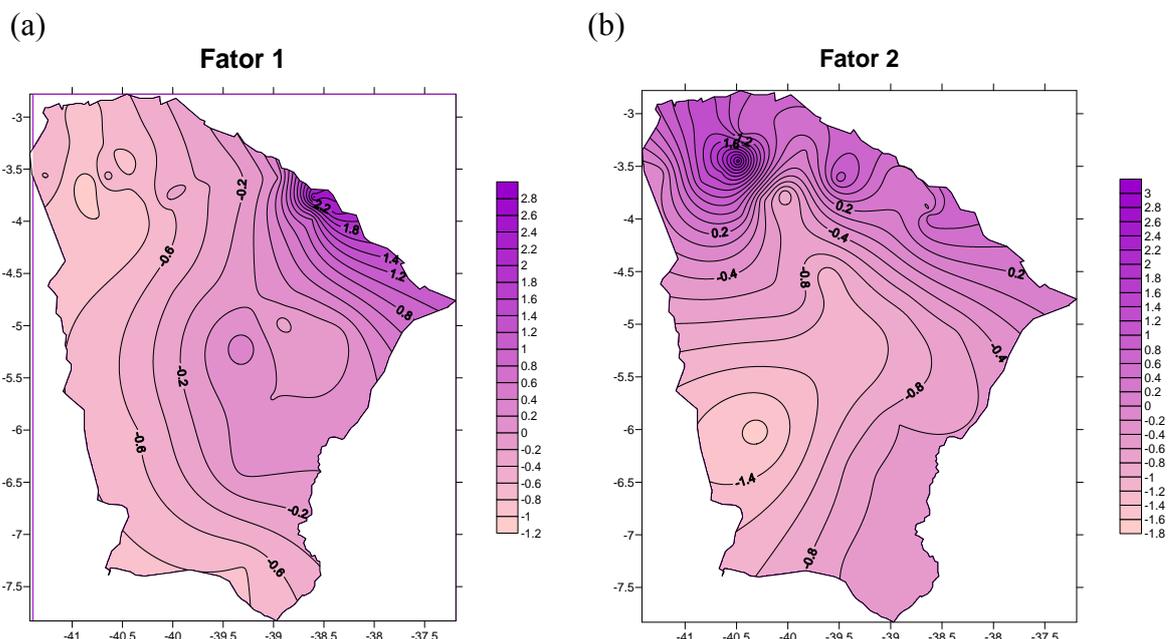


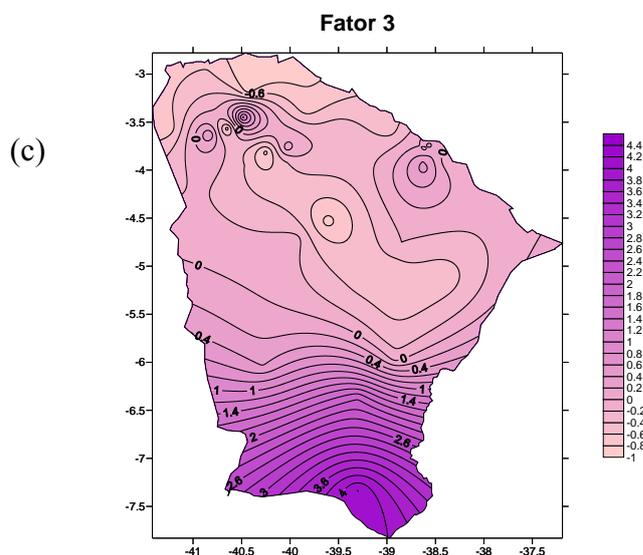
**Figura 3:** Cargas fatoriais rotacionadas (correlações) para os três fatores comuns de precipitação que explicam 93,37% do total (55,20%+24,51%+13,56%) no estado do Ceará.

O primeiro fator comum temporal rotacionado da precipitação, que explicou 55,20 % da variância total da série, apresenta correlação significativa no intervalo de 0,8 a nos meses de junho a setembro e entre 0,6 e 0,8 em maio (Figura 3). Observa-se pela distribuição espacial deste fator, que as maiores contribuições (pesos), com valores superiores a 2, encontram-se na região nordeste do Ceará e valores inferiores a 1 no restante do estado (Figura 4a).

O segundo fator comum que explicou 24,51% da variância total dos dados de precipitação pluvial tem alta correlação com valores entre 0,8 e 1 nos meses de fevereiro a abril (Figura 3). O padrão espacial associado a este fator (Figura 4b) apresenta valores superiores a 2,0 na região noroeste do Ceará. Neste período, a distribuição de chuvas (não mostrada) está associada à atuação mais intensa da ZCIT (Uvo, 1989).

O terceiro fator comum explicando 13,56% da variância apresenta altas correlações entre 0,8 e 1 nos meses de outubro a dezembro, porém com correlações significativas entre 0,6 e 0,8 em janeiro e maio (Figura 3). A configuração espacial deste fator apresenta as maiores contribuições da precipitação com valores superiores a 2,0 na região sul e centro sul do Ceará (Figura 4c). A distribuição deste último fator relaciona às chuvas de outubro a dezembro (não mostrada) com a entrada de frentes e de VCAN (Kousky, 1979).





**Figura 4** - Padrões espaciais dos fatores comuns dos totais médios mensais da precipitação pluvial referente aos fatores no estado do Ceará para o período de 1970-2000: (a) fator1, (b) fator2 e (c) fator3.

#### 4. CONCLUSÕES

A Análise Fatorial das séries temporais representada pelos padrões de precipitação possibilitaram uma melhor compreensão dos aspectos físicos responsáveis pela variabilidade sazonal do parâmetro estudado. As três primeiras componentes explicaram em torno de 93% da variância total dos dados.

O peso do primeiro fator comum mostrou que as maiores contribuições ocorreram na região metropolitana de Fortaleza. Do segundo fator, apresentou maiores contribuições na região do noroeste cearense. E por fim, o peso do terceiro fator evidenciou o regime pluviométrico na parte do centro sul e cariri cearense. Portanto, para explicar a variabilidade espacial da precipitação sobre o Ceará é necessário apenas as três primeiras componentes principais.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

JOHNSON, R.A.; WICHERN, D.W. **Applied Multivariate Statistical Analysis**. 4 ed. Upper Saddle River: Prentice Hall, 1998.

KOUSKY, V. E. Frontal influences on Northeast Brazil. *Monthly Weather Review*, 107 (9): 1140 - 1153, 1979.

PEARSON, K. On lines and planes of closest fit to systems of points in space. *Phil. Mag.*, v. 6, p. 559-72, 1901.

SILVA, J. B.; CUNHA, F. B.; NETO, W. P. G. Modelagem das Chuvas trimestrais por regiões homogêneas no Estado do Ceará. **Revista Brasileira de Agrometeorologia – Santa Maria**, v.9, n. 2, p. 317-324, 2001 b.

SILVA, A. M.; COSTA; D. L. C. R.; LINS, C. J. C. Precipitações no Nordeste Brasileiro: tendências de variação e possíveis implicações na agricultura. **V Semana do Meio Ambiente**. 3 a 5 de junho de 2008 – Recife, PE.

UVO, C.B. A Zona de Convergência Intertropical (ZCIT) e sua relação com a Precipitação na Região Norte do Nordeste Brasileiro. **Dissertação de Mestrado em Meteorologia**, INPE-São José dos Campos, 1989.