

# PREVISÃO DE TOTAL ANUAL DE PRECIPITAÇÃO

POR

J.M.Damazio<sup>1</sup>, J.Kelman<sup>1,2</sup>

RESUMO-- Apresenta-se uma sequência de procedimentos para escolher entre modelos alternativos de previsão de total anual de precipitação. Esta sequência é aplicada na série de Fortaleza resultando na seleção de um modelo que atribui ao período 1982/86 probabilidades mais elevadas de ocorrência de baixas precipitações do que em média se poderia esperar.

## INTRODUÇÃO

É crescente no Brasil o interesse na utilização de modelos de previsão de secas com a antecipação de um a cinco anos, destacando-se os estudos feitos com a série de precipitação anual de Fortaleza. Teixeira et al (1980) apresentam uma revisão sobre o assunto, e corroboram a hipótese de que haverá um período de estiagem no período de 1979/86 para o Nordeste. Com relação a este prognóstico, pode-se colocar duas questões:

- a) Total anual de precipitação é um bom índice do fenômeno seca ?  
- Sem dúvida a resposta seria positiva caso reservatórios de acumulação estivessem disponíveis na área de interesse. Na sua falta, a simples previsão do total anual é insuficiente, sob o ponto de vista da agricultura, uma vez que essa total pode ocorrer em poucos dias.

---

<sup>1</sup>Pesquisador do Centro de Pesquisas de Energia Elétrica - CEPTEL.

<sup>2</sup>Professor Visitante da Coordenação do Programa de Pós-Graduação em Engenharia - COPPE/UFRJ.

- b) A série de Fortaleza é representativa do que ocorre no Nordeste?  
- Os coeficientes de correlação entre a série de Fortaleza e demais séries da região se situam entre 0,5 e 0,8 (Teixeira et al (1980)). Mesmo que se pudesse prever perfeitamente a precipitação em Fortaleza, restaria ainda um grande grau de incerteza nas extrapolações feitas para outras localidades.

Neste artigo a série de Fortaleza é utilizada para ilustrar um conjunto de procedimentos destinados a discriminar entre modelos de previsão competitivos. A metodologia utilizada pode ser empregada em outros contextos, tais como, para verificar a viabilidade de predição de estiagens para o sistema hidroelétrico brasileiro. Neste caso, precipitação anual é seguramente um índice relevante devida à grande capacidade de regularização dos reservatórios do sistema.

#### OS MODELOS PROPOSTOS PARA A SÉRIE DE PRECIPITAÇÃO ANUAL EM FORTALEZA

A tabela 1 relaciona 130 totais anuais (outubro a setembro) de precipitação em Fortaleza, 1848/49 a 1977/78, fornecida por Strang (1979). Os valores de assimetria (0,34) e curtose (2,77) destes dados sugerem o uso da distribuição normal. O bom ajuste desta distribuição aos dados é confirmado pela estatística qui-quadrado (0,85) obtida com sete intervalos equiprováveis. Usando quatro graus de liberdade este valor tem o nível de confiança de 94%. Portanto, pode-se construir o primeiro modelo para representar a série de precipitações anuais em Fortaleza:

$$M1: y_t = \mu + a_{1,t} \quad (1)$$

onde  $y_t$  é a precipitação anual no ano  $t$ ,  $\mu$  é a média de população e  $a_{1,t}$  o ruído normal, independente, média zero e variância  $\sigma_1^2$ . Convém observar que de acordo com a equação (1) os valores de precipitação de anos diferentes são independentes entre si.

A figura 1 mostra o periodograma da série de Fortaleza. Nota-se a existência de dois picos, correspondendo aos ciclos de 13 e 26 anos, que tem induzido diversos pesquisadores a adotar a hipótese de ciclicidade para esta série (Teixeira et al (1980)).

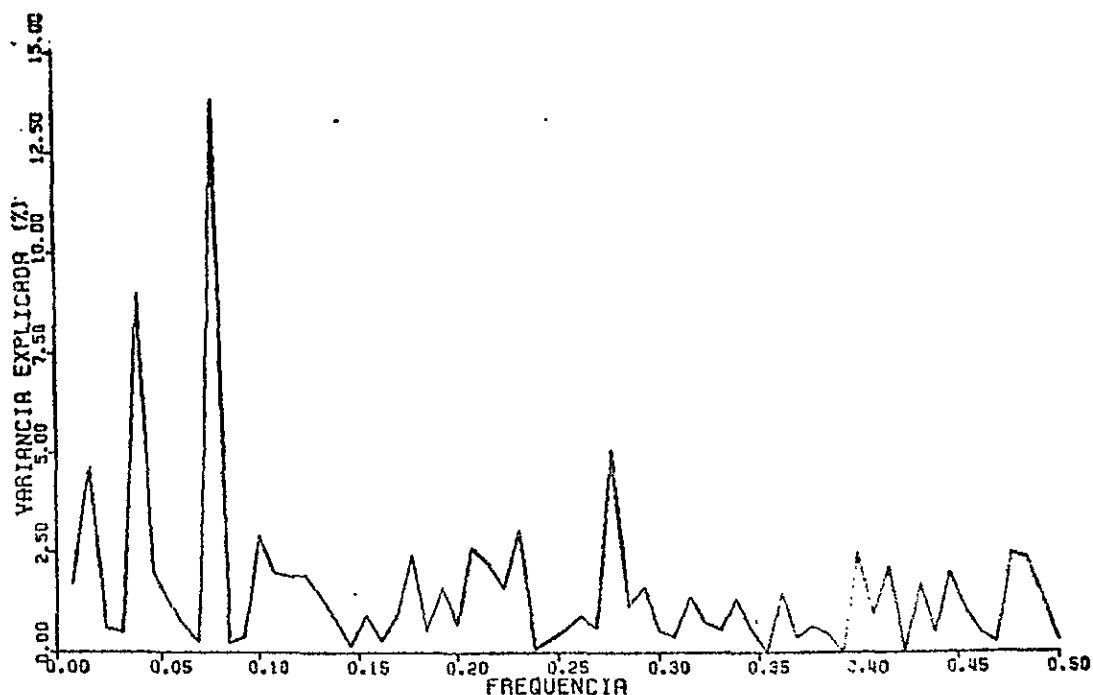


Figura 1 - Periodograma da série de 130 anos de precipitação anual em Fortaleza.

Fisher (1929) determinou a distribuição de probabilidades da máxima ordenada do periodograma debaixo da hipótese nula de normalidade e independência, correspondente ao modelo M1. Na série de Fortaleza o pico do periodograma ocorre no ciclo de 13 anos com um valor de 13,7%, que tem o baixíssimo nível de significância de 0,5%. Ou seja, pode-se rejeitar o modelo M1 em favor do modelo:

$$M2: y_t = \mu + A \cos\left(\frac{2\pi}{13}t + B\right) + a_{2,t} \quad (2)$$

onde A e B são respectivamente amplitude e fase correspondentes ao ciclo de 13 anos e  $a_{2,t}$  o ruído normal, independente, média zero e variância  $\sigma_2^2$ .

A significância do ciclo de 26 anos pode ser avaliada repletando-se o teste de Fisher com os resíduos  $a_{2,t}$  da equação (2) (Yevjevich (1972)). O harmônico de 26 anos explica 10,4% da variância residual, o que corresponde ao nível de significância de 5,7%. Considerar ou não significativo o ciclo de 26 anos é uma questão subjetiva. De qualquer forma o modelo M3 será considerado como uma das alternativas a serem comparadas no próximo item:

$$M3: y_t = \mu + A \cos\left(\frac{2\pi}{13} t + B\right) + C \cos\left(\frac{2\pi}{26} t + D\right) + a_{3,t} \quad (3)$$

onde A e B são os mesmos da equação (2), C e D são respectivamente amplitude e fase correspondente ao ciclo de 26 anos e  $a_{3,t}$  o ruído normal, independente, média zero e variância  $\sigma_3^2$ .

A rejeição do modelo M1 pode conduzir a formulação dos modelos cíclicos M2 e M3, ou, alternativamente, ao uso de modelos estocásticos com memória, por exemplo, do tipo autorregressivo. O mais simples destes modelos é o chamado modelo markoviano, cuja equação é:

$$M4: y_t = \mu + \rho(y_{t-1} - \mu) + a_{4,t} \quad (4)$$

onde  $\rho$  é a autocorrelação lag-um e  $a_{4,t}$  novamente ruído normal, independente, média zero e variância  $\sigma_4^2$ . A figura 2 mostra o periodograma dos ruídos em (4) estimados por mínimos quadrados com os dados de Fortaleza. Nota-se a permanência dos picos correspondentes aos ciclos de 13 e 26 anos. No entanto, a aplicação do teste de Fisher, neste caso, resulta no nível de significância de 22% para o ciclo de 13 anos. Ou seja, não se pode rejeitar a hipótese de que a série de ruídos estimados seja independente serialmente e portanto que M4 seja correto.

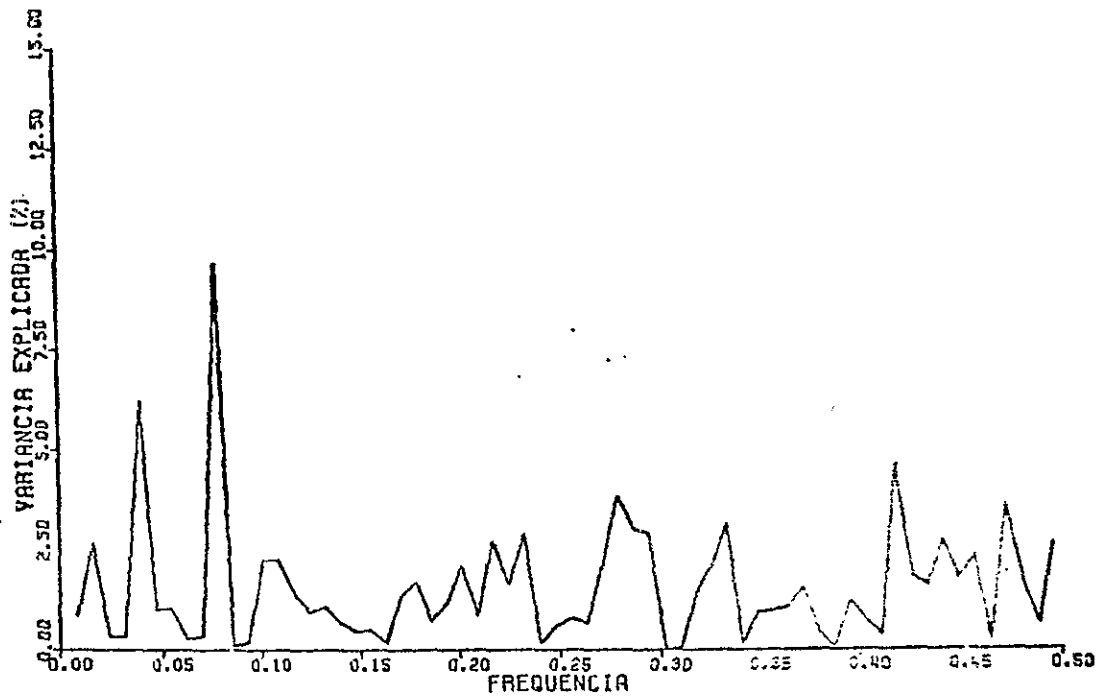


Figura 2 - Periodograma dos Ruídos do Modelo Markoviano Aplicado aos 130 anos de Precipitação anual em Fortaleza.

Na tabela 2 apresenta-se as estimativas de mínimos quadrados dos parâmetros de cada modelo, as correspondentes estimativas das variâncias residuais, e as estatísticas de portmentau com correspondentes níveis de significância.

MODELO	PARÂMETROS	$\sigma_k^2$	$Q_k$ $m=30$	GRAUS DE LIBERDADE	%Q's > $Q_k$
1	$\hat{\mu}=1427$	238020	36,7	30	18
2	$\hat{A}= 256,47$ $\hat{B}= 1,42$	205142	28,4	28	5
3	$\hat{C}= 206,88$ $\hat{D}= 1,58$	183742	17,0	26	62
4	$\hat{\rho}= 0,23$	225438	20,1	29	89

Tabela 2 - Estimação de Parâmetros

A estatística de portamanteau é indicada para a verificação da hipótese de independência do ruído sendo calculada por:

$$Q_k = N \sum_{i=1}^m r_i^2(a_k) \quad (5)$$

onde  $r_i(a_k)$  é a correlação lag  $i$  dos ruídos do modelo  $k$ , e  $m$  usualmente maior que 15 (usou-se  $m=30$  para os cálculos da tabela 2). Para o cálculo do nível de significância usa-se a distribuição  $\chi^2(m-n_k+1)$ , sendo  $n_k$  o número de parâmetros do modelo  $k$ . Pelos resultados da tabela 2 nenhum dos modelos é rejeitado ao nível de confiança de 10%.

#### A ESCOLHA DO MELHOR MODELO

Markham (1974) justifica os ciclos de 13 e 26 anos como resultados da "ressonância de um longo período no sistema oceano-atmosfera, governando a posição média da linha de convergência intertropical, bem como as temperaturas oceânicas equatoriais". Pode-se argumentar que a influência da posição da linha de convergência intertropical na precipitação induza memória no processo, mas não necessariamente ciclicidade (Dias (1981)). O fato é que o conhecimento físico do processo é ainda insuficiente para que se possa prescindir de métodos estatísticos para escolha de um modelo de previsão. Dada a série de 130 valores da tabela 1,  $\{y_1, \dots, y_{130}\}$  e os quatro modelos propostos no item anterior, o método estatístico deve encontrar o modelo mais apropriado para representar  $\{y_1, \dots, y_{130}\}$ .

## O critério de Akaike

De acordo com Akaike (1974) o uso de testes de hipóteses na identificação de modelos é inapropriado já que hipóteses nulas são aproximações e portanto diferentes da realidade. Esta contradição só pode ser resolvida formulando explicitamente o problema de identificação como um problema de aproximação da realidade mais complexa. A regra de Akaike consiste em considerar o modelo que fornece o menor valor da estatística AIC como aquele que melhor se aproxima do modelo real.

$$AIC_k = N \ln \sigma_k^2 + 2n_k \quad (6)$$

A tabela 3 mostra os valores obtidos com a série de precipitação anual em Fortaleza.

MODELO	AIC <sub>k</sub>
M1	1611
M2	1596
M3	1586
M4	1606

Tabela 3 - Critério de Akaike

Por este critério os modelos são ordenados por ordem de preferência, em M3, M2, M4, M1.

## Um critério por amostras particionadas

Um critério mais pragmatico para a escolha da representação matemática a ser usada nas previsões a longo prazo consiste em usar parte da amostra para a estimação de parâmetros deixando o conjunto mais recente de observações para se testar as previsões de cada modelo. Aqui se adotará um critério de máxima verossimilhança através dos seguintes procedimentos:

- i) Estimação dos parâmetros dos modelos usando os 104 primeiros valores da série de observações.
- ii) Determinação da verossimilhança,  $f_k$ ,  $k=1, 4$  das 26 últimas observações debaixo da hipótese de cada modelo e parâmetros estimados no item anterior.
- iii) A probabilidade de cada modelo será calculada por:

$$P[M_k] = \frac{f_k}{\sum_j f_j} \quad (7)$$

- iv) Escolher o modelo de maior probabilidade.

Na tabela 4 apresenta-se os resultados.

MODELO	Parâmetros Estimados Com $y_1, \dots, y_{104}$	$\sigma_k^2$ Estimado Com $y_1, \dots, y_{104}$	$P[M_k   y_{105}, \dots, y_{130}]$
1	$\hat{\mu} = 1400$	231407	0,00
2	$\hat{A} = 238,91$ $\hat{B} = 1,50$	202869	0,10
3	$\hat{C} = 174,05$ $\hat{D} = 1,39$	187722	0,90
4	$\hat{\rho} = 0,21$	221202	0,00

Tabela 4 - Critérios de Amostras Particionadas

É de se observar que tanto o critério de Akaike quanto o critério de amostras particionadas apontam para o modelo M3, ficando em segundo lugar o modelo M2.



## PREVISÃO DE ANOS SECOS EM FORTALEZA

Todos os quatro modelos propostos para o cálculo de previsões implicam em se considerar as precipitações anuais,  $y_t$ , como o resultado da adição de um componente aleatório,  $a_{x,t}$ , a um componente determinístico que pode ser constante (M1) ou variável no tempo (M2, M3) ou ainda, determinado por condições passadas (M4). É usual referir-se à componente determinística como "a previsão", ficando a variância do componente aleatório responsável pelo cálculo de intervalos de confiança para a precipitação. Outra utilização, mais interessante e completa, dos modelos consiste em se considerar a distribuição de probabilidades do componente aleatório (normal para os modelos considerados) para o cálculo da probabilidade de ocorrência de ano seco. Como exemplo classifica-se arbitrariamente como ano seco em Fortaleza aquela cujo total de precipitação for menor que 1017mm ( $P[Y_t < 1017 | M] = 0,20$ ). A tabela 5 mostra as "previsões" e as probabilidades de ocorrência de ano seco em Fortaleza fornecidas pelo modelo M3 (recomendado no item anterior) para os próximos cinco anos. Vale notar que pelo modelo M1 a previsão e a probabilidade de ocorrência de ano seco são constantes e respectivamente iguais a 1472mm e 0,20. Na tabela 6 compara-se as probabilidades de ocorrência de 0, 1, 2, 3, 4 e 5 anos secos em Fortaleza nos próximos cinco anos fornecidos pelos modelos M1 e M3.

ANO	E(Yt) "PREVISÃO" (mm)	$P[Y_t < 1017]$
81/82	1005	0,51
82/83	1035	0,48
83/84	1122	0,40
84/85	1244	0,30
85/86	1372	0,20

Tabela 5 - Previsão e Probabilidades de Ano seco em Fortaleza, Modelo M3.

x		0	1	2	3	4	5
P [ X=x ]	M1	0,328	0,410	0,205	0,051	0,006	0,000
	M3	0,086	0,283	0,356	0,211	0,058	0,006

x - Número de anos secos

Tabela 6 - Distribuição de Probabilidades do Número de Anos Secos em Fortaleza no período entre 1981/82 e 1985/86.

### CONCLUSÕES

Os dois critérios de discriminação aplicados indicam a conveniência de se aceitar tanto o harmônico de 13 anos como de 26 anos para prever totais anuais de precipitação em Fortaleza.

As tabelas 5 e 6 demonstram que a consideração destes harmônicos modifica sensivelmente as previsões quanto à ocorrência de baixas precipitações em Fortaleza em relação aos valores obtidos pela hipótese de independência. Por exemplo, de acordo com o modelo recomendado (M3) a probabilidade de ocorrência de algum "ano seco" em Fortaleza no período 81/82 - 85/86 é de 0,91, portanto maior do que a probabilidade que em média se poderia esperar em um período de cinco anos (0,67). "Ano seco" foi arbitrado como aquele que ocorre em média uma vez em cada cinco anos.

ANO	TOTAL	ANO	TOTAL	ANO	TOTAL
1848/49	2.044	1898/99	2.434	1948/49	1.875
1849/50	838	1899/00	746	1949/50	1.083
1850/51	1.750	1900/01	1.733	1950/51	692
1851/52	1.362	1901/02	870	1951/52	1.463
1852/53	1.256	1902/03	797	1952/53	1.067
1853/54	1.634	1903/04	1.142	1953/54	1.032
1854/55	1.196	1904/05	1.196	1954/55	1.055
1855/56	1.734	1905/06	1.427	1955/56	905
1856/57	1.848	1906/07	695	1956/57	1.209
1857/58	1.292	1907/08	833	1957/58	519
1858/59	1.527	1908/09	907	1958/59	1.488
1859/60	1.646	1909/10	2.022	1959/60	992
1860/61	1.277	1910/11	1.488	1960/61	1.697
1861/62	1.679	1911/12	2.477	1961/62	1.310
1862/63	1.437	1912/13	1.870	1962/63	1.883
1863/64	1.129	1913/14	1.542	1963/64	2.653
1864/65	1.233	1914/15	508	1964/65	1.636
1865/66	2.453	1915/16	1.230	1965/66	1.271
1866/67	864	1916/17	2.108	1966/67	1.856
1867/68	1.263	1917/18	1.373	1967/68	1.348
1868/69	1.498	1918/19	683	1968/69	1.848
1869/70	1.586	1919/20	1.768	1969/70	1.139
1870/71	1.332	1920/21	2.481	1970/71	2.112
1871/72	2.343	1921/22	1.623	1971/72	1.293
1872/73	1.992	1922/23	1.580	1972/73	2.339
1873/74	1.592	1923/24	1.758	1973/74	2.427
1874/75	1.485	1924/25	1.212	1974/75	1.806
1875/76	1.693	1925/26	1.584	1975/76	1.477
1876/77	483	1926/27	1.164	1976/77	1.873
1877/78	511	1927/28	1.013	1977/78	1.670
1878/79	570	1928/29	1.240		
1879/80	1.512	1929/30	1.080		
1880/81	1.461	1930/31	1.167		
1881/82	1.247	1931/32	882		
1882/83	1.431	1932/33	914		
1883/84	1.007	1933/34	1.745		
1884/85	1.327	1934/35	1.809		
1885/86	1.399	1935/36	829		
1886/87	1.342	1936/37	1.300		
1887/88	725	1937/38	1.570		
1888/89	765	1938/39	1.908		
1889/90	1.521	1939/40	1.459		
1890/91	1.097	1940/41	925		
1891/92	1.017	1941/42	781		
1892/93	1.532	1942/43	1.006		
1893/94	2.533	1943/44	1.120		
1894/95	2.504	1944/45	1.758		
1895/96	2.200	1945/46	1.648		
1896/97	1.827	1946/47	1.597		
1897/98	859	1947/48	1.611		

Tabela 1 - Totais Anuais de Precipitações (mm)  
em Fortaleza.

## REFERÊNCIAS

- AKAIKE, H. (1974), "A new look at Statistical Model Identification", Institute of Electrical and Electronics Engineers Transaction on Automatic Control, Volume AC-19, number 6, pags. 716-723.
- DIAS, P.L. (1981), Comunicação Pessoal.
- FISHER, R.A. (1929), "Tests of significance in harmonic analysis", Proceedings of Royal Society of London, série A, volume 125, pags. 54-59.
- MARKHAM, C.G. (1974), "Apparent Periodicities in Rainfall at Fortaleza, Ceará, Brazil", Journal of Applied Meteorology, volume 13, pags. 176-179.
- STRANG, D.M.G.D. (1979), Utilização dos dados pluviométricos de Fortaleza, CE, visando determinar probabilidades de anos secos e chuvosos, Relatório Técnico ECA-03/79, Instituto de Atividades Espaciais, Centro Técnico Aeroespacial, 12200 - São José dos Campos - São Paulo - Brasil.
- TEIXEIRA, L., GIRARDI, C., GUEDES, R.L. (1980), Resumo da análises sobre a série pluviométrica de Fortaleza-Ceará - Brasil, Relatório Técnico ECA-02/80, Instituto de Atividade Espaciais, Centro Técnico Aereo-espacial, 12200, São José dos Campos - São Paulo, Brasil.
- YEVJEVICH, V. (1972), "Structural analysis of hydrologic time series", Hydrology Papers, number 56, Colorado State University, Fort Collins, Colorado 80521.

## ABSTRACT

It is presented a methodology to choose among models which were proposed to forecast total anual precipitation. The methodology is applied to the data of Fortaleza and results in the selection of a cyclic model which attributes to the period 1982/86 a high probability of low precipitation.