Revista Brasileira de Geofísica; 1987, *Vol. 5*, 53-62

doi: 10.22564/rbgf.v5i1.1064

PREDICTION OF ANNUAL RAINFALL AT FORTALEZA, CE, BRAZIL IN RECENT YEARS

R.P. KANE

Instituto de Pesquisas Espaciais – INPE
C.P. 515, 12201 – São José dos Campos, SP, Brasil

Power Spectrum Analysis of Fortaleza annual rainfall data for 1849-1976 indicated periodicities T = 12.9 and 25.1 years with amplitude r_k significant at a $5\sigma_{rk}$ (a priori) level and in addition, T = 2.07, 3.63, 4.84, 5.69, 10.1, 18.0 and 61.0 at a $2\sigma_{rk}$ (a priori) level. Using these periodicities, the predicted values for 1979-86 matched well with the observed values. Predictions for the period 1987-2015 include a minor drought during 1993-96 and a major drought during 2003-12. However, the above periodicities explain only ~62% Variance, implying ~38% random component. The standard errors for predicted flood and drought rainfalls are $\sigma = \pm 400$ mm and $\sigma = \pm 300$ mm, respectively. Hence, drought intervals could have a few years of normal (or even above-normal) rainfall. Similarly, individual drought years could occur during expected flood intervals. A transient QBO (T = 2-3 years) could complicate matters further.

A análise do Espectro de Potência dos dados de precipitação anual de Fortaleza, para o período que vai de 1849 a 1976, indica periodicidades T = 12,9 e 25,1 anos, com amplitudes r_k significantes em um nível de $5\sigma_{rk}$ (a priori) e, além disso, periodicidades T = 2,07, 3,63, 4,84, 5,69, 10,1, 18,0 em um nível de $2\sigma_{rk}$ (a priori). Usando todas estas periodicidades, os valores previstos para 1977 até 1986 correspondem bem aos valores observados. Para 1987-2015, são previstas uma seca menor no período de 1993 a 1996 e uma seca maior no período de 2003 a 2012. Todavia, as periodicidades acima explicam somente ~62% da Variância, o que implica em ~38% de componente aleatória. Os desvios padrão para as precipitações de inundação e seca previstas são $\sigma = \pm 400$ mm e $\sigma = \pm 300$ mm, respectivamente. Por isso, os intervalos de seca poderiam ter alguns anos de precipitações normais (ou até acima do normal). Do mesmo modo, poderiam ocorrer alguns anos de secas durante os intervalos de inundação esperados. Um QBO transiente (T = 2 a 3 anos) poderia complicar as coisas ainda mais.

INTRODUCTION

Northeast Brazil suffers frequently from severe droughts. The longest rainfall data series available in that region is for Fortaleza, Ceara (mean 1430 mm), starting from 1849.

Using the annual rainfall values for 12 years at Fortaleza, Markham (1974) conducted a spectral analysis by the Blackman and Tukey (1958) method and reported significant periodicities at T = 13 and 26 years, but did not attempt any predicitions. Girardi & Teixeira (1978) selected these periodicities, attributed them *arbitrary* peak-to-trough ranges of 1300 mm and 2000 mm, respectively, and adjusted their phases to match some of the earlier drought years. Extrapolating for later years, they predicted a prolonged drought during 1979-83. This prediction seems to have come true. The average annual rainfall for 36 locations in Northeast Brazil for 1979, 1980, 1981, 1982 and 1983 deviated -10%, -15%, -17%, -18% and –44%, respectively, from the general average (Kane, 1985). Using more rigorous methods of analysis, Kane & Trivedi (1984, 1986a) noticed that the amplitudes of the two major peaks at T = 12.9 and 25.1 years were only ~250 mm for each (peak-to-trough 500 mm); but the predictions of drought for 1979-83 were essentially correct. They also predicted recovery from 1984 (inclusive) onwards and for future, mild droughts during 1993-95 and a major drought during 2004-10.

In this communication, the validity of predictions for the last 3 years (1984, 1985, 1986) will be examined and some limitations of this methodology will be discussed. It may also be mentioned beforehand that the rainfall regimes in different parts of Northeast Brazil are not alike. The monthly rainfall reaches maximum in the northern part (Ceara) during

March-April, while for the southern part (Bahia) the maximum occurs in November-December and for the eastern coastal areas it occurs in May-July (Kousky, 1979). Also, apart from the fact that some parts receive as little as 400 mm of annual rainfall while some others receive as much as 2000 mm, the year-to-year variations are not similar and sometimes, the northern and southern parts show opposite extremes of rainfall (droughts versus floods). A correlation analysis shows (Kane & Trivedi, 1986b) that the similarity with Fortaleza (correlation coefficient + 0.60) extends only to about 600 km from Fortaleza.

METHOD OF ANALYSIS

The most sophisticated method of power spectrum analysis now available is MESA (Maximum Entropy Spectral Analysis) developed by Burg (1967) and critically reviewed by Ulrych and Bishop (1975), who show the superiority of MESA over the earlier Blackman & Tukey (1958) method. Using several geophysical parameters as well as artificial samples, Kane (1977, 1979) confirmed the superiority of MESA. However, MESA has two major defects. Firstly, though MESA has the advantage that it can detect periodicities as large as the data length, the larger periodicities can be wrong by as much as 20%. If analysis is confined to periodicities about 50% of the data length (e.g. up to T = 50 years in 100 years data), this error is not large (less than 2%). Secondly, the amplitude estimates in MESA are unreliable (Kane & Trivedi, 1982). Hence Kane (1977) suggested an alternative way, viz. to use MESA only for detecting the periodicities $T_1, T_2, \ldots, T_n$ and then use these in the expression:

$$f(t) = A_0 + \sum_{k=1}^{n} \left(a_k \operatorname{Sin} 2\pi \frac{t}{T_k} + \right.$$

$$\left. + b_k \operatorname{Cos} 2\pi \frac{t}{T_k}\right) + E$$

$$= A_0 + \sum_{k=1}^{n} r_k \sin\left(2\pi \frac{t}{T_k} + \Phi_k\right) + E,$$

where f(t) = the observed time series and E is the error factor. Using Multiple Regression Analysis (Johnston, 1960; Bevington, 1969), the best estimates of A_0, (a_k, b_k) can be obtained (by a least-square fit) and their standard errors estimated. From these, the amplitudes r_k and their standard errors σ_{rk} can also be calculated. The Percentage Variance Explained (PVE) by every r_k is given by $50\ (r_k^2/\sigma^2)$, where σ^2 = Variance of the time series f(t).

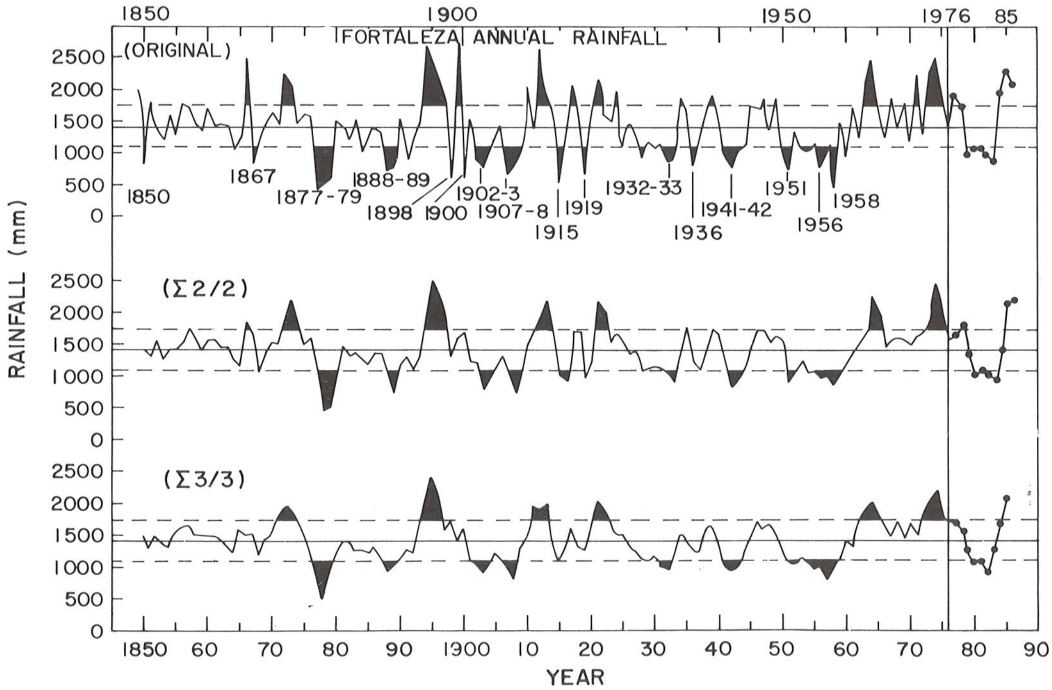

Figure 1 – Fortaleza annual rainfall series for 1849-1976 (full lines) and for 1977-1986 (big dots). Top curve represents original values and the second and the third rows show moving averages over two and three successive values, respectively. Excess and deficit rainfall (high and low) are shown shaded black. Years of well-known droughts are indicated.

ANALYSIS

Using Fortaleza data for 1849-1976 (128 years), the 16 periodicities detected by MESA and their amplitudes obtained by a Multiple Regression Analysis are shown in Table 1 for Group 1. Some of these are not significant even at a $2\sigma_{rk}$ (a priori) level. In Group 2, we included only the 9 periodicities significant at a $2\sigma_{rk}$ (a priori) level and repreated the Multiple Regression Analysis. The amplitudes now obtained are slightly different from those in Group 1. For Group 3, only the 5 periodicities significant at a $3\sigma_{rk}$ (a priori) level in Group 1 were selected. Finally, Group 4 included only the two periodicities T = 12.9 and 25.1 years, which are significant at $5\sigma_{rk}$ (a priori) level as seen in Group 1.

For each group separately, the parameters A_0, (a_k, b_k) were estimated and the expected values were calculated by using the expression:

$$\overline{f}(t) = A_0 + \sum_{k=1}^{n} (a_k \sin 2\pi \frac{t}{Tk} + b_k \cos 2\pi \frac{t}{Tk})$$

not only for 1849-1976, but for later years too. For 1977 onwards, the expected values could thus be compared with independent data.

Fig. 1 is a plot of the Fortaleza values for 1849 onwards, original values at the top and moving averages over 2 successive values in the middle plot and 3 successive values in the bottom plot.

Fig. 2 shows a plot of the observed values (solid lines) and the expected values (crosses and dashes) for 1944 onwards only. The vertical line marks 1976. Observed values for 1977 onwards are shown by bigger dots. As can be seen, the period 1979-83 is indicated as a period of drought in all four groups and the observed values confirm this (values below average). For 1984, 1985, 1986, Groups 1, 2 and 3 show positive expected deviations which have also been verified. For Group 4, the expected recovery was slower than observed.

Table 1 – Amplitudes (mm) and their standard errors (mm) for the 16 periodicities T = 2.07, 2.24, 2.37, 2.77, 3.02, 3.39, 3.63, 4.42, 4.48, 5.69, 8.71, 10.1, 12.9, 18.0, 25.1, 61.0 years (designated as Group 1) detected by MESA for Fortaleza annual rainfall series for 1849-1976 (128 years, series mean 1420 mm). Groups 2, 3, 4 include only those periodicities significant at 2σ, 3σ and 4σ levels (a priori), respectively. The right half gives the Percentage Variance Explained by the various periodicities.

Period T (years)	Group 1	Group 2	Group 3	Group 4	Percentage variance explained (group 1)			
					All Peaks	Peaks significant at an *a priori*		
						2σ Level	3σ Level	4σ Level
2.07	146	140	138		4.3	4.3	4.3	
2.24	81				1.3			
2.37	72				1.1			
2.77	81				1.3			
3.02	46				0.4			
3.39	96				1.8			
3.63	139	150			3.9	3.9		
4.42	88				1.6			
4.84	125	116			3.2	3.2		
5.69	125	130			3.2	3.2		
8.71	94				1.8			
10.1	147	143	143		4.4	4.4	4.4	
12.9	249	251	253	248	13.6	13.6	13.6	13.6
18.0	105	101			2.2	2.2		
25.1	243	245	246	243	12.0	12.0	12.0	12.0
61.0	166	169	164		5.8	5.8	5.8	
Standard Error	±48	±48	±52	±56	61.9 Total	52.6 Total	40.1 Total	25.6 Total

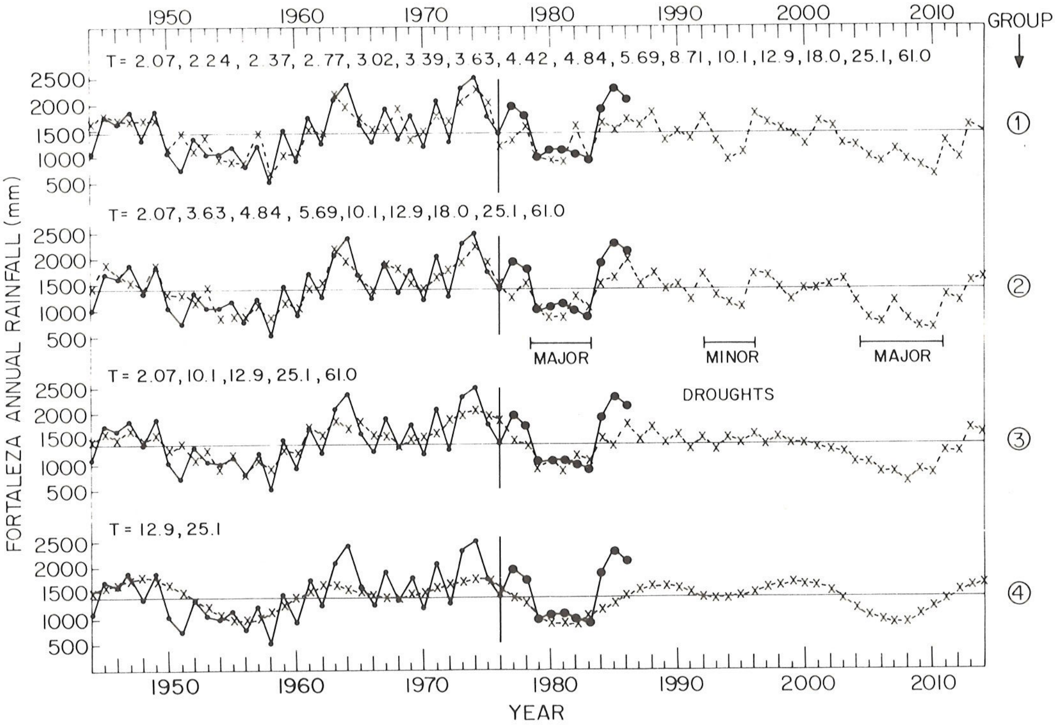

Figure 2 – Fortaleza rainfall, original values for 1944-76 (full lines) and for 1976-86 (full lines and big dots). The crosses and dashed lines show the predicted values for Groups 1-4, using periodicities T (years) as indicated.

Fig. 3 shows a plot of observed values versus expected values. Values for 1977-86 are shown as big dots. Except for 1977, for which the expected value was below average while the observed value was far above average, and for 1982, for which the expected value was above average while the observed value was below average, the observed values for other recent years (1977-86) match well with the expected values. The overall correlation coefficient is +0.59, which is not very high. This is mainly because of the large scatter, for which there are two reasons. Firstly, as seen from Table 1, all the 16 periodicities of Group 1 together explain only a maximum possible 61.9% of the Variance of the series. Thus, at least 38% remains unexplained and should be attributed to a substantial random component. Secondly, some of the periodicities, specially the QBO (Quasi-Biennial Oscillation, T = 2-3 years), seem to be of a very transient nature. In Fig. 1, QBO seems to be very prominent during some intervals (e.g., 1956-72) and negligible in some others (e.g., 1979-83). Thus, the estimated QBO amplitudes in Table 1 represent only an average of a widely varying parameter. This variability would add to the scatter. The expected values were divided into 6 groups of about 21 years each of (i) severe droughts, (ii) mild droughts, (iii) small negative deviations, (iv) small positive deviations, (v) mild floods and (vi) severe floods. The *averages* of observed values for the extreme groups are indicated in Fig. 3 by triangles with their standard error bars (full line), as well as the extreme values (dashed extensions). Thus, for positive extremes (floods, expected mean value 2000 mm), the mean observed value was 2104±400 mm with extremes between 1385 and 2793. For negative extremes (droughts, mean expected value 900 mm), the mean observed value was 1022±300 mm with extremes between 503 and 1575 mm. Thus, a very large scatter is involved, yielding an uncertainty $\sigma = \pm 300$ mm in the drought estimates and $\sigma = \pm 400$ mm in the flood estimates. It should be expected, therefore, that some of the expected drought (or flood) years could turn out to be near normal rainfall years. Normally, one would not expect an expected drought year to turn out to be a flood year or vice versa. However, Fig. 1 shows that, in the past, flood (drought) years have followed drought (flood) years and also that floods or droughts have occurred in consecutive years. Table 2 gives details of these occurrences during 1849-1982.

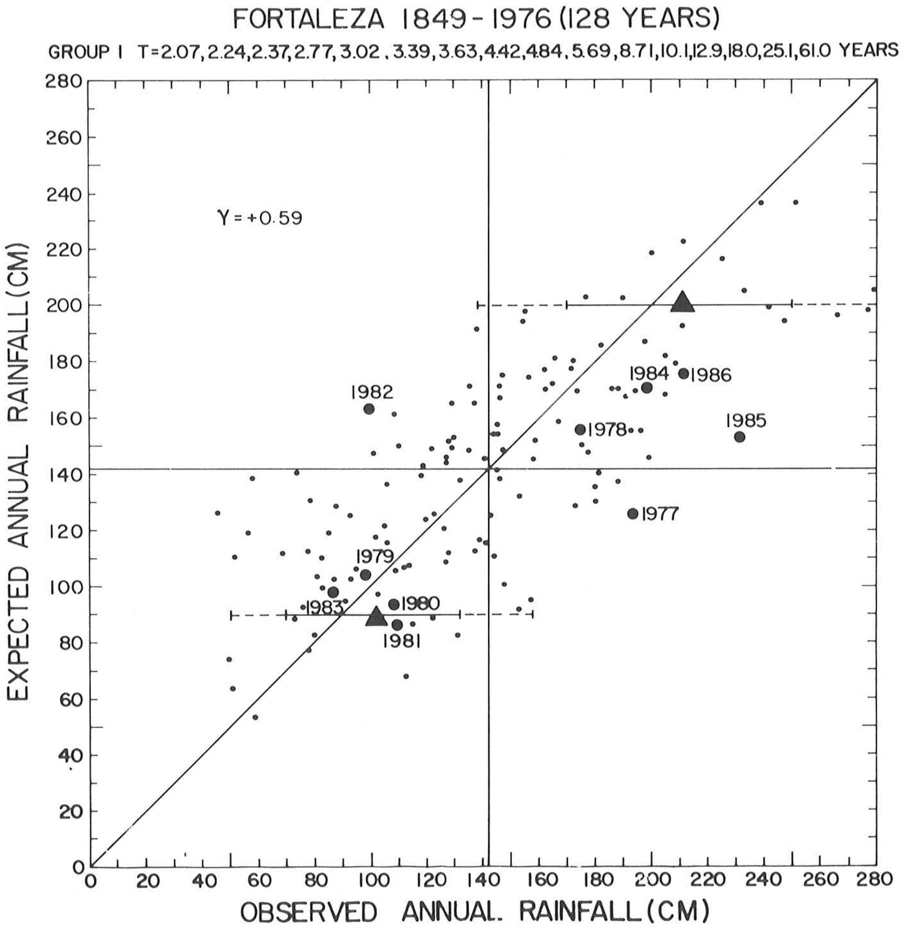

Figure 3 – Fortaleza annual rainfall values, observed versus expected. Big dots represent recent years (1977-86) as marked. The triangles with standard errors (horizontal full lines) and extreme values (dashed extensions) represent the mean observed values for expected extreme rainfall. The expected values were obtained using all the 16 periodicities T of Group 1.

In Table 1, we gave details about the significance of the various periodicities. However, since the QBO is seen to be transient or non-stationery, it is necessary to check the stationarity of the other periodicities. For this, the data were divided into two equal parts, each of 65 years (1849-1913, 1914-1978) and MESA and Multiple Regression Analysis were carried out for each part separately. The amplitudes (mm) of some significant periodicities were as follows:

Periodicity years (Approximate)	2.07	2.37	3.0	3.4	4.9	9.8	13	18	20-26	Std error
1849-1913	185	216	156	182	178	227	269	186	306	65
1914-1978	190	130	160	200	128	130	254	180	250	55

Table 2 – Characteristics of Fortaleza annual rainfall series (1849-1982).

Type	Normal years	Low rainfall (Drought) years	High (excess) Rainfall years	Years of contrast (Low followed by high, or vice versa)
Single years	1890 1955 1901 1957 1911 1959 1916 1968 1920 1970 1935 1972 1940 1976 1948 1978 1952 (17 years) *Variance* (1%)	1884 1936 1891 1956 1915 1958 1919 1960 1928 (9 years) *Variance* (12%)	1856 1949 1921 1967 1924 1969 1934 1971 1939 1977 1947 (11 years) *Variance* (8%)	
Two consecutive years	1892-93 1922-23 1937-38 1961-62 1965-66 (10 years) *Variance* (1%)	1888-89 1902-03 1907-08 1932-33 1950-51 1953-54 (12 years) *Variance* (12%)	1872-73 1917-18 1963-64 (6 years) *Variance* (11%)	1866 (High), 1867 (Low) 1909 (Low), 1910 (High) (4 years) *Variance* (7%)
Three or more consecutive years	1852-55 (4 years) 1857-65 (9 years) 1868-71 (4 years) 1874-76 (3 years) 1880-83 (4 years) 1885-87 (3 years) 1904-06 (3 years) 1925-27 (3 years) 1929-31 (3 years) 1944-46 (3 years) (39 years) *Variance* (5%)	1877-79 (3 years) 1941-43 (3 years) 1979-82 (4 years) (10 years) *Variance* (12%)	1894-97 (4 years) 1912-14 (3 years) 1973-75 (3 years) (10 years) *Variance* (23%)	1849 (High), 1850 (Low), 1851 (High) 1898 (Low), 1899 (High), 1900 (Low) (6 years) *Variance* (8%)
Total	(66 years) *Variance* (7%)	(31 years) *Variance* (36%)	(27 years) *Variance* (42%)	(10 years) (5 High, 5 Low) *Variance* (15%)

Total number of years (1849-1982) = 134 years
Total rainfall for 134 years + 190659 mm
Total variance for 134 years + 32994191 $(mm)^2$
Normal years = Rainfall 1423 ∓ 331 = ~1090 to 1760 mm
Low rainfall (drought) years = Rainfall less than 1090 mm
High (excess) rainfall years = Rainfall more than 1760 mm
Year = Calendar year (January-December)
Mean rainfall (1849-1982) = 1423 mm
Mean variance (1849-1982) = 246225 $(mm)^2$
Standard devidation σ = 496 mm
Probable error = 0.66 σ = 331 mm

As can be seen, not only the QBO (2-3 years) but some of the other periodicities, notably T = 4.9 and 9.8 years, have different amplitudes in the two halves. Hence, this time series seems to be partly *nonstationery*. In Fig. 2, Group 1 includes all periodicities and the predictions for the same are highly unreliable. For Group 2, 3 and 4, periodicities used have larger and larger degree of significance and hence the reliability of predictions may be somewhat better. Periodicities of Group 4 (specially T = 13 years) seem to be stationery. However, Group 4 accounts for only 25% Variance. Hence no good reliability is expected for any group. The large scatter in Figure 3 is indicative of this unreliability.

Fig. 4, 5, 6 show plots similar to Fig. 3, but for Groups 2, 3, 4. For Group 2 (Fig. 4), out of the 10 big dots for 1977-86, only one (1977) is in the wrong quadrant (expected value low, observed value high). The overall correlation is lower (+0.49), indicating a larger scatter as compared to Group 1 (Fig. 3), mainly because Group 2 has only 9 periodicities, explaining at the most 52.6% of the total Variance, leaving at least 47% for the random component. But the prediction is better than that in Group 1 in the sense that, out of the 10 observed values for 1977-86, 8 were in proper quadrants (++ or - -) for Group 1, while 9 were in proper quadrant for Group 2.

Group 3 (Fig. 5) contains only 5 periodicities, explaining at the most 40.1% of the total Variance. Thus, at least 60% goes to the random component (correlation only +0.37). The values for 1978 is in the wrong quadrant. Thus, Group 3 does not show any better performance than Group 2.

Group 4 (Fig. 6) contains only the two most prominent periodicities T = 12.9 and 25.1 years, explaining at the most 25.6% of the total Variance. The scatter is large, correlation very small (+0.23) and the matching for the last few years (1984-86) is not satisfactory, mainly because the observed values recovered much faster than expected. Thus, utilizing only the two most prominent periodicities did not give good predictions. Additional (secondary) periodicities are necessary for better predictions.

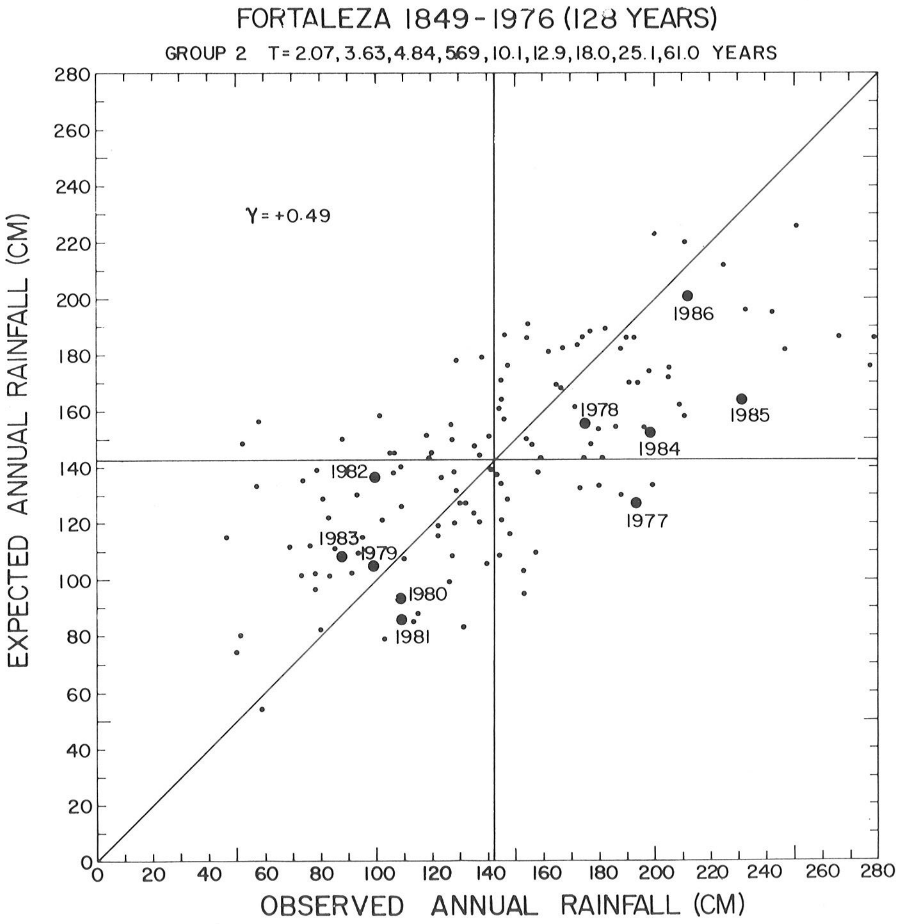

Figure 4 – Same as Fig. 3, for Group 2, with 9 periodicities T, as indicated.

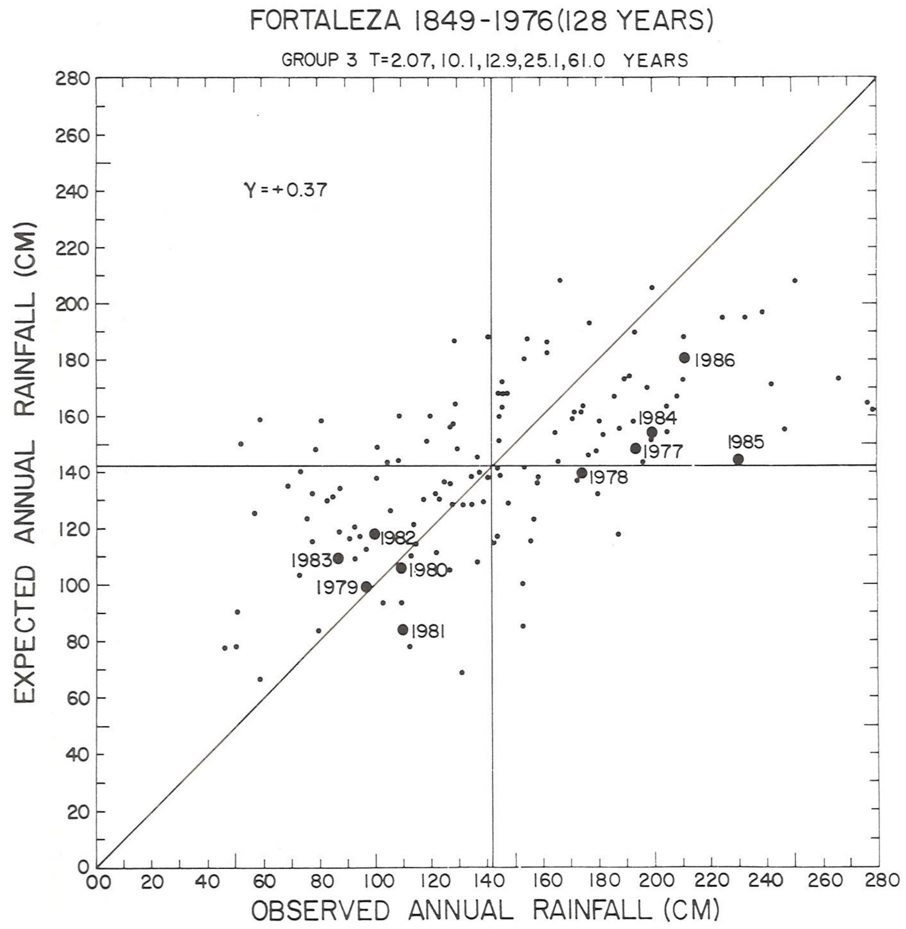

Figure 5 – Same as Fig. 3, for Group 3, with 5 periodicities T, as indicated.

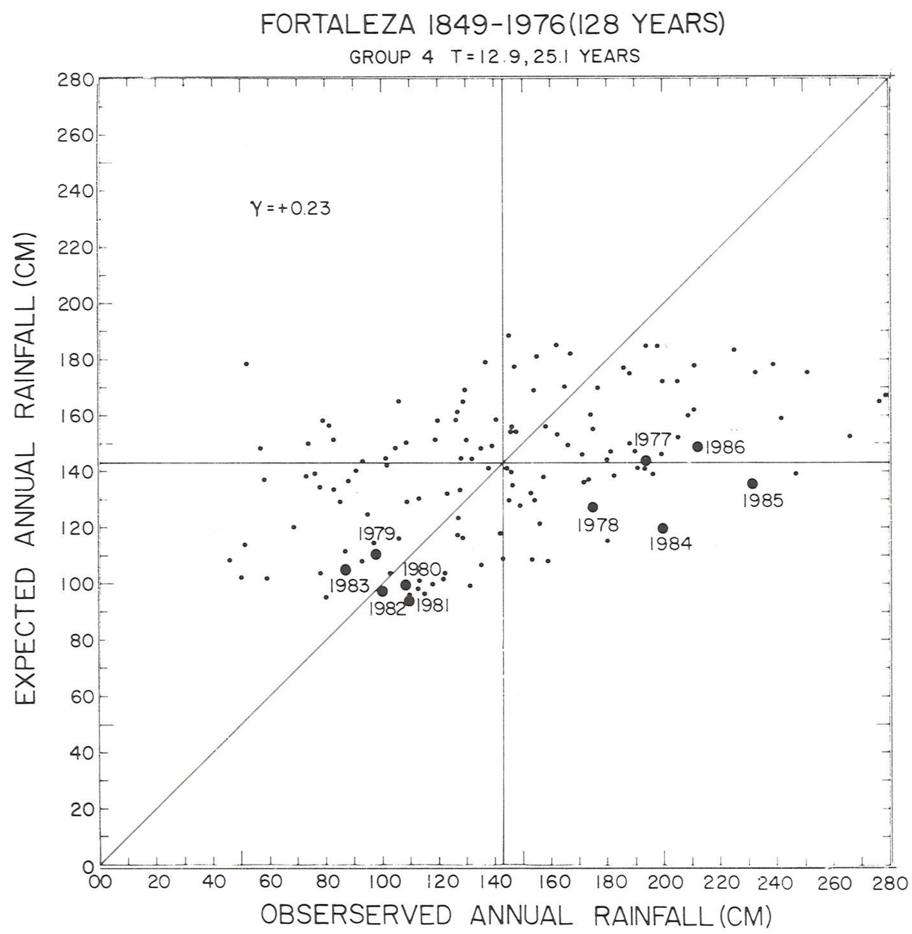

Figure 6 – Same as Fig. 3, for Group 4, with only 2 periodicities T = 12.9 and 25.1 years.

Table 3 – Observed and expected values for 1977-86 and expected values for future years (unit: mm) for Fortaleza annual rainfall (mean ~1430 mm).

Year	Observed	EXPECTED			
		(16 Peaks) Group 1	(9Peaks) Group 2	(5 Peaks) Group 3	(2 Peaks) Group 4
1977	1941	1260	1274	1489	1433
1978	1752	1554	1559	1399	1261
1979	985	1039	1054	984	1104
1980	1095	926	934	1075	991
1981	1100	855	852	836	942
1982	1004	1637	1366	1183	966
1983	884	978	1099	1095	1057
1984	1981	1703	1522	1549	1194
1985	~2350	1521	1648	1431	1351
1986	2140	1748	2029	1794	1498
1987		1630	1524	1533	1608
1988		1860	1717	1756	1665
1989		1332	1435	1413	1667
1990		1492	1495	1591	1623
1991		1378	1218	1301	1554
1992		1738	1665	1516	1483
1993		1324	1357	1329	1433
1994		910	1158	1554	1422
1995		1063	1088	1421	1454
1996		1804	1740	1579	1520
1997		1661	1706	1457	1603
1998		1497	1427	1522	1678
1999		1415	1247	1426	1720
2000		1276	1413	1419	1710
2001		1647	1460	1371	1641
2002		1550	1482	1288	1519
2003		1142	1583	1265	1362
2004		1208	1154	1079	1196
2005		877	824	1068	1054
2006		870	752	812	963
2007		1098	1196	884	939
2008		933	801	675	988
2009		817	689	937	1100
2010		601	664	859	1252
2011		1295	1307	1305	1417
2012		943	1165	1284	1563
2013		1623	1591	1746	1666
2014		1492	1635	1601	1712
2015		2229	1928	1897	1701

Table 3 lists the observed values and values expected by using periodicities in Groups 1, 2, 3 and 4 for 1977-86 (upper part) and the expected values for future years. From these and the plots in Fig. 2, there are indications of a minor drought during 1993-96 and a major drought during 2003-12. However, due to a large random component (40%) and due to the transient nature, especially of the QBO, a few individual years in these intervals may have near normal or perhaps even above normal rainfall. But the *average* rainfall (average for the few years) during these intervals is indicated to be below normal. As such, precautionary measures could be taken and planned, starting from now, to be executed within the next decade or two.

Since some periodicities are likely to be transient, this analysis needs to be updated as and when fresh data are available. In the present analysis, we used data only for 1849-1976, leaving 1977-86 for checking the prediction. In Kane & Trivedi (1984), the results for 1849-1983 were given and *confirmed* the above predictions of droughts for 1993-96 and 2003-12. Recently, data for 1984-86 were also added and the analysis repeated. Forecasts of drought for 1993-96 (minor) and 2003-12 (major) were confirmed.

CONCLUSION

(1) Power spectrum analysis of the Fortaleza annual rainfall data for 1849-1976 (128 years) showed two very prominent periodicities (significance $5\sigma_{rk}$, a priori) at T = 12.9 and 25.1 years. Additional periodicities T = 2.07, 10.1 and 61.0 years were significant at a $3\sigma_{rk}$ (a priori) level. Further periodicities T = 3.63, 4.84, 5.69 and 18.0 years were significant at a $2\sigma_{rk}$ (a priori) level. For good predictions, not only the two most significant periodicities but several others significant at a $2\sigma_{rk}$ (a priori) level were needed. For the last 10 years (1977-86), the expected and observed values were in the same sense (both above average or both below average) for 9 years.

(2) There is a large random component (~40%) in the Fortaleza annual rainfall data (average about 1430 mm). As such, predicted values can have standard errors as large as 400 mm in above-average rainfalls and 300 mm in the below-average rainfalls. Hence, occasionaly drought years during predicted flood years or vice versa are not ruled out. Subject to this limitation, a mild drought during 1993-96 and a major drought during 2003-12 are envisaged.

ACKNOWLEDGEMENTS

This work was partially supported by FNDCT Brazil under contract FINEP 537/CT.

REFERENCES

BEVINGTON, P.R. – 1969 – "Data reduction and error analysis for the Physical Sciences". McGraw-Hill Book Co. New York pp. 164-176.

BLACKMAN, R.B. & TUKEY, J.W. – 1958 – "The measurements of Power Spectra", Dover, New York.

BURG, J.P. – 1967 – Maximum Entropy Spectral Analysis, Paper presented at the 37th Meeting, Soc. of Explor. Geophys., Oklahom City.

GIRARDI, C. & TEIXEIRA, L. – 1978 – Prognóstico do tempo a longo prazo, Instituto de Atividades Espaciais (IAE), Centro Técnico Aerosespacial (CTA), São José dos Campos, SP (Brasil). Relatório Técnico ECA-06/78, December.

JOHNSTON, J. – 1960 – "Econometric methods". Mc-Graw-Hill Book Co. New York pp. 134-135.

KANE, R.P. – 1977 – Power Spectrum Analysis of Solar and Geophysical Parameters. J. Geomag. Geoelect. **29**:471-495.

KANE, R.P. – 1979 – Maximum entropy Spectral Analysis of some artificial samples. J. Geophys. Res. **84**:965-966.

KANE, R.P. – 1985 – An appraisal of the duration of the recent drought in northeast Brazil. Rev. Bras. Geof. **3**:115-123.

KANE, R.P. & TRIVEDI, N.B. – 1982 – Comparison of Maximum Entropy Spectral Analysis (MESA) and Least-Squares Linear Prediction (LSLP) methods for some artificial samples. Geophysics **47**:1731-1736.

KANE, R.P. & TRIVEDI, N.B. – 1984 – Prediction of drought in the Brazilian Northeast region from an analysis of periodicities in the Fortaleza rainfall data, INPE-3334-RPI/118, November.

KANE, R.P. & TRIVEDI, N.B. – 1986a – Are droughts predictable? Climatic Change **8**:209-223.

KANE, R.P. & TRIVEDI, N.B. – 1986b – Comparison of rainfall characteristics at Fortaleza (Ceara) and other nearby locations in Northeast Brazil. INPE-3851-PRE/915, April.

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

KOUSKY, V.E. & CHU, P.S. – 1978 – Fluctuations in annual rainfall for Northeast Brazil. J. Meteor. Soc., Japan **56**:457-465.

MARKHAM, C.G. – 1974 – Apparent periodicities in rainfall at Fortaleza, CE, Brazil. J. Applied Meteorology **13**:176-179.

ULRYCH, T.J. & BISHOP, T.N. – 1975 – Maximum Entropy Spectral Analysis and Autoregressive Decomposition. Rev. Geophys. **13**:183-200.

Versão original recebida em Abr./87
Versão final, em Abr./88