

ANÁLISE HARMÔNICA ¹

EDILBERTO AMARAL²

Sinopse

Os fundamentos teóricos da análise harmônica e a marcha da análise são apresentados. Um exemplo é estudado, no qual, embora os dados originais não se distribuam normalmente, a variância não sendo independente da média, os contrastes necessários à análise harmônica apresentam variâncias homogêneas, dispensando prévia transformação dos dados (precipitações mensais em Pelotas, Rio Grande do Sul, 1900/51). Em outro exemplo, a transformação é estudada e realizada, e os dados transformados são analisados (precipitações mensais em Morro Velho, Minas Gerais, 1855/1951).

A estabilidade das componentes harmônicas das precipitações mensais em Pelotas é testada pelo método da análise da variação, revelando-se significativas as ondas semestral, quadrimestral e anual, as duas primeiras mostrando-se predominantes no curso anual das precipitações mensais (máximos em fevereiro e em setembro). O teste de completicidade indica que 91,2% da variação das precipitações médias mensais corre à conta dessas três ondas.

Não se dispondo dos dados anuais de Morro Velho, não foi possível submetê-los ao teste de estabilidade. Quase toda a variação entre as médias mensais (99,3%) corre à conta da onda anual, como é típico em um clima tropical continental (chuvas convectivas no verão, seca hiberna).

Como contribuições originais, foi estabelecida a distribuição da diferença entre os quadrados de duas amplitudes, introduziu-se o conceito de coeficiente de correlação cíclica, que não deve ser confundido com o conhecido coeficiente de correlação serial, e adaptou-se o método de Fuhrich, permitindo sua aplicação na análise harmônica, o que tem importância prática nos casos em que as componentes harmônicas são em número considerável, determinando-se, então, apenas aquelas componentes cujas amplitudes são relevantes.

Inclui em apêndice as deduções matemáticas, a fim de aliviar o texto. Apresenta também alguns parágrafos sobre testes de normalidade e transformação de dados, visando à homogeneização das variâncias.

INTRODUÇÃO

O objetivo do presente trabalho é eminentemente didático. Do ponto de vista técnico, deve entender-se como trabalho preliminar, pretendendo-se aplicar o

método na análise harmônica de médias ou totais de elementos climáticos, na base de pântadas (períodos de cinco dias).

Métodos mais gerais de análise periodográfica, aplicáveis quando o período fundamental não é conhecido, são apenas mencionados, devendo constituir assunto de um futuro trabalho.

O presente trabalho foi remodelado e ampliado, após ter sido apresentado no Curso de Pós-Graduação em Climatologia Agrícola, realizado na Escola Superior de Agricultura "Luiz de Queiroz" Piracicaba, São Paulo, de julho a setembro de 1965, sob o patrocínio do Serviço de Meteorologia do Ministério da Agricultura.

O estudo sobre a distribuição da diferença entre os quadrados de duas amplitudes, sobre seu significado e a adaptação já referida do método de Fuhrich, constituem, possivelmente, contribuições originais. Esta última contribuição não só parece constituir uma natural introdução ao método de Fuhrich de análise periodográfica, como também permite determinar as componentes na ordem decrescente de suas amplitudes, o que tem importância prática quando o número de componentes harmônicas é considerável.

¹ Recebido para publicação em 25 de maio de 1966. Boletim Técnico n.º 56 do Instituto de Pesquisas e Experimentação Agropecuárias do Sul (IPEAS), Caixa Postal E, Pelotas, Rio Grande do Sul.

² Eng.º Agrônomo da Seção de Documentação e Estatística do IPEAS e Professor Catedrático de Matemática da Escola de Agronomia Eliseu Maciel, Pelotas, Rio Grande do Sul.

GENERALIDADES

A análise harmônica é um dos métodos de estudo dos fenômenos periódicos ou quase periódicos³. Desses métodos, é ela, sem dúvida, o mais bem sistematizado, sendo de aplicação obrigatória quando o período fundamental é conhecido e o processo não se reduz, sabidamente, a uma onda senoidal simples, tal como ocorre com o curso anual ou diurno de um elemento climático em uma localidade.

Uma função $x=x(t)$, onde a variável independente t não representa, obrigatoriamente, tempo, diz-se periódica de período T se

$$x(t + T) \equiv x(t),$$

isto é, se $x(t + T)$ é idêntico a $x(t)$ ou, o que é o mesmo, se $x(x + T) = x(t)$ qualquer que seja o valor de t . É o que ocorre, por exemplo, com a função

$$x = a \text{ sen } (ft + A),$$

função senoidal simples, que é periódica de período $T = 2\pi/f$, se f e A são expressos em radianos, e de período $T = 360/f$, se f e A são expressos em graus. Quando t varia de $-\infty$ a $+\infty$, $\text{sen}(ft + A)$ assume, periodicamente, todos os valores compreendidos entre -1 e $+1$ e, portanto, $x = a \text{ sen } (ft + A)$ assume todos os valores compreendidos entre $-a$ e $+a$. A constante a representa, assim, a metade da oscilação da função⁴ e se denomina *amplitude* da oscilação; $f=2\pi/T$ radianos, ou $f=360/T$ graus, é a frequência angular e $f/2\pi$, ou $f/360$, é a frequência isto é, o número de oscilações completas por unidade da variável t (por unidade de tempo, se é esse o caso). O ângulo A , em radianos ou em graus, é o *ângulo fase*; duas funções senoidais

$$x_1 = a_1 \text{ sen } (ft + A_1),$$

$$x_2 = a_2 \text{ sen } (ft + A_2),$$

que têm o mesmo período (e, portanto, a mesma frequência), mas diferentes ângulos fase, estão *defasadas*, sendo $A_2 - A_1$ a diferença de fase, expressa em graus ou radianos; pode-se escrever

$$x_2 = a_2 \text{ sen } \left[f \left(t + \frac{A_2 - A_1}{f} \right) + A_1 \right],$$

³ Por fenômeno quase periódico, entende-se, com Brooks e Carruthers (1953), aquele em que o valor do elemento estudado não se reproduz exatamente em certos intervalos de tempo, como se a um processo estritamente periódico, se superpusessem influências modificadoras, de caráter aleatório.

⁴ Por oscilação de uma função, entende-se a diferença não negativa entre seu extremo superior e seu extremo inferior.

de modo que, se $A_2 > A_1$ e t é o tempo, a onda senoidal x_1 está atrasada em relação a x_2 de $(A_2 - A_1)/f$ unidades de tempo.

Num circuito de corrente alternada puramente resistivo, a tensão e a intensidade da corrente estão *em fase*, isto é, ambas se anulam simultaneamente e passam simultaneamente pelos valores críticos correspondentes (máximos e mínimos), como se pode ver na Fig. 1, segundo Van. Valkenburgh *et al.* (1960).

Num circuito puramente indutivo, a tensão e a corrente estão defasadas de 90° (Fig. 2), isto é, de um quarto do período.



FIG. 1. Circuito resistivo.

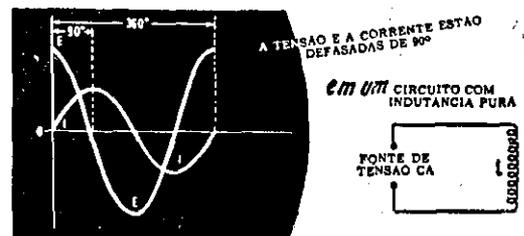


FIG. 2. Circuito indutivo.

Se o circuito tem resistência e reatância indutiva, a onda da corrente (intensidade) está atrasada em relação à tensão de um ângulo compreendido entre zero e 90° ; assim, por exemplo, se a resistência e a reatância indutiva são iguais, a defasagem é de 45° (Fig. 3).

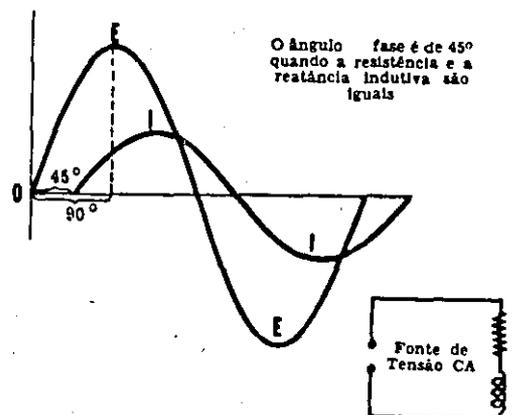


FIG. 3. Circuito resistivo-indutivo.

As Figs. 2 e 3 são da mesma procedência da Fig. 1.

Na Fig. 4, segundo Blackwood *et al.* (1962), representa-se um instrumento usado para estudar os sons. Fazendo vibrar um diapasão junto ao bocal do diafragma, este será impellido para dentro e para fora pelas ondas sonoras, transmitindo-se o movimento alternativo ao ponteiro e traçando-se uma curva ondulante numa película que se desloca com velocidade constante.

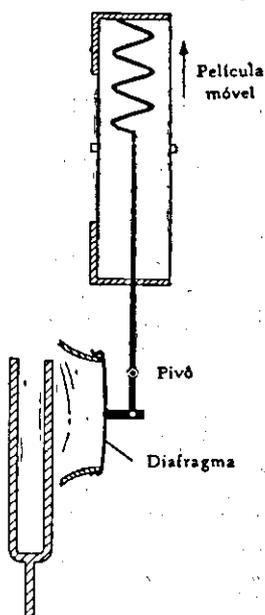


FIG. 4. Instrumento para estudo dos sons.

A Fig. 5 mostra as curvas produzidas por diapasões de frequência (A), de 100 vibrações por segundo, (B), de 200 vibrações por segundo, (C), de 600 vibrações por segundo e (D), pelos três diapasões vibrando simultaneamente. Como se pode ver, comparando com a Fig. 6, o som produzido pelos três

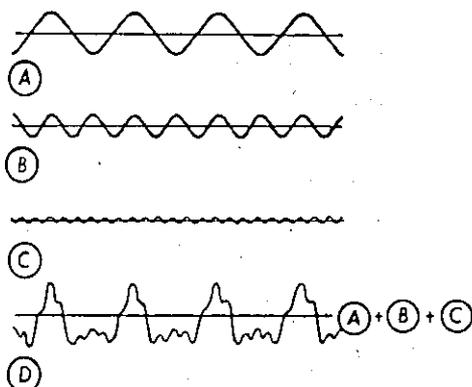


FIG. 5. Traços devidos a diapasões de diferentes frequências.

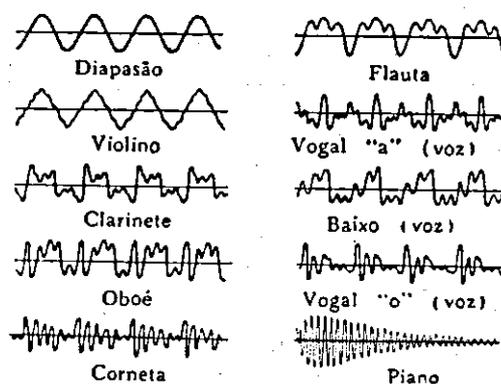


FIG. 6. Traços produzidos por sons de vários instrumentos e da voz humana.

diapasões, vibrando simultaneamente, assemelha-se ao de um clarinete.

“Quando dois sons soam bem juntos, eles estão em harmonia. O grande matemático e filósofo Pitágoras descobriu a lei da harmonia há vinte e cinco séculos. Dois sons estão em harmonia se a relação entre suas frequências for igual à razão entre dois números inteiros pequenos”. (Blackwood *et al.* 1962).

O som de menor frequência, num grupo de sons em harmonia, é o primeiro harmônico ou som fundamental; o de frequência n vezes maior, sendo n inteiro, é o harmônico de ordem n . Assim, na Fig. 5, são ilustrados os sons produzidos pelo primeiro, pelo segundo e pelo sexto harmônico. A curva representada em (D), na mesma figura, resulta, pois, da síntese dos três harmônicos. A comparação com a curva produzida pelo clarinete (Fig. 6) sugere a possibilidade de realizar a análise harmônica, isto é, discriminar os sons puros que, soando em conjunto, reproduzem o som do clarinete.

Outros exemplos de fenômenos periódicos são a altura do Sol acima do horizonte ao meio dia em uma dada latitude, e a hora do nascer e a do pôr do Sol, também em uma latitude dada.

Muitos fenômenos meteorológicos, astronômicos e geofísicos, como o fenômeno das marés, o número de manchas solares, o curso diurno e o anual de elementos climáticos, da temperatura, por exemplo, são fenômenos quase periódicos. O tema do presente trabalho é a análise harmônica desses fenômenos.

FUNDAMENTOS TEÓRICOS

Demonstra-se (Ver apêndice, itens de 1 a 7), que, dividindo-se a circunferência em n arcos iguais, $r = 2\pi/n$ radianos, ou $r = 360/n$ graus, valem as identidades.

$$\begin{array}{rcl}
 \cos 0 + \cos r & + \cos 2r & + \dots + \cos (n-1)r = 0 \\
 \cos 0 + \cos 2r & + \cos 4r & + \dots + \cos 2(n-1)r = 0 \\
 \dots & \dots & \dots \\
 \cos 0 + \cos I_n r & + \cos 2I_n r & + \dots + \cos I_n(n-1)r = 0 \\
 \sin 0 + \sin r & + \sin 2r & + \dots + \sin (n-1)r = 0 \\
 \sin 0 + \sin 2r & + \sin 4r & + \dots + \sin 2(n-1)r = 0 \\
 \sin 0 + \sin 3r & + \sin 6r & + \dots + \sin 3(n-1)r = 0 \\
 \sin 0 + \sin I_{n-1} r & + \sin 2I_{n-1} r & + \dots + \sin I_{n-1}(n-1)r = 0 \\
 \dots & \dots & \dots
 \end{array}$$

onde $I_n = I(n/2)$ é o maior inteiro que não supera $n/2$ e I_{n-1} é o maior que não supera $(n-1)/2$.

Verificam-se, ainda, as identidades:

$$\begin{array}{rcl}
 \cos 0 \cos 0 & + \cos ir \cos jr & + \cos 2ir \cos 2jr & + \dots + \cos (n-1)ir \cos (n-(n-1)jr = 0 \\
 \sin 0 \sin 0 & + \sin ir \sin jr & + \sin 2ir \sin 2jr & + \dots + \sin (n-1)ir \sin (n-(n-1)jr = 0 \\
 \sin 0 \cos 0 & + \sin ir \cos jr & + \sin 2ir \cos 2jr & + \dots + \sin (n-1)ir \cos (n-(n-1)jr = 0
 \end{array}$$

sendo i e j inteiros positivos, $i \neq j$; e, também,

$$\begin{array}{l}
 \cos^2 0 + \cos^2 ir + \cos^2 2ir + \dots + \cos^2 (n-1)ir = n/2 \text{ (n ímpar ou } i < n/2, \text{ se n é par).} \\
 \cos^2 0 + \cos^2 ir + \cos^2 2ir + \dots + \cos^2 (n-1)ir = n \text{ (} i = n/2, \text{ n par).} \\
 \sin^2 0 + \sin^2 ir + \sin^2 2ir + \dots + \sin^2 (n-1)ir = n/2 \text{ (n ímpar ou } i < n/2, \text{ se n é par),}
 \end{array}$$

Resulta das fórmulas precedentes ortogonalidade dos $n-1$ contrastes*

$$\begin{array}{rcl}
 x_1 \cos 0 + x_2 \cos r & + x_3 \cos 2r & + \dots + x_n \cos (n-1)r \\
 x_1 \cos 0 + x_2 \cos 2r & + x_3 \cos 4r & + \dots + x_n \cos 2(n-1)r \\
 x_1 \cos 0 + x_2 \cos 3r & + x_3 \cos 6r & + \dots + x_n \cos 3(n-1)r \\
 \dots & \dots & \dots \\
 x_1 \cos 0 + x_2 \cos I_n r & + x_3 \cos 2I_n r & + \dots + x_n \cos I_n(n-1)r \\
 x_1 \sin 0 + x_2 \sin r & + x_3 \sin 2r & + \dots + x_n \sin (n-1)r \\
 x_1 \sin 0 + x_2 \sin 2r & + x_3 \sin 4r & + \dots + x_n \sin 2(n-1)r \\
 x_1 \sin 0 + x_2 \sin 3r & + x_3 \sin 6r & + \dots + x_n \sin 3(n-1)r \\
 \dots & \dots & \dots \\
 x_1 \sin 0 + x_2 \sin I_{n-1} r & + x_3 \sin 2I_{n-1} r & + \dots + x_n \sin I_{n-1}(n-1)r,
 \end{array}$$

onde $r = 2\pi/n$ radianos ou $r = 360/n$ graus, $I_n = I(n/2)$ é o maior inteiro que não supera $n/2$ e I_{n-1} é o maior inteiro que não supera $(n-1)/2$.

Se n é ímpar, $I_{n-1} = I_n = (n-1)/2$ e a norma⁶ de tôdas as $n-1$ formas lineares é igual a $n/2$.

Se n é par, $I_n = n/2$ e $I_{n-1} = (n/2) - 1$; neste caso, a norma da forma linear

$$x_1 \cos 0 + x_2 \cos I_n r + x_3 \cos 2I_n r + \dots + x_n \cos I_n(n-1)r$$

é igual a n e a de tôdas as demais é igual a $n/2$.

Demonstra-se que, dados $n-1$ contrastes ortogonais entre n números, x_1, x_2, \dots, x_n , contrastes que serão designados por $X_1, X_2, \dots,$

X_{n-1} , a soma dos quadrados das diferenças entre os n números e sua média aritmética pode exprimir-se na forma

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \frac{X_1^2}{N_1} + \frac{X_2^2}{N_2} + \dots + \frac{X_{n-1}^2}{N_{n-1}},$$

onde \bar{x} é a média aritmética dos x_i , e N_1, N_2, \dots, N_{n-1} são normas dos contrastes respectivos. Dêsse modo, o conjunto dos $n-1$ graus de liberdade entre os n números, x_1, x_2, \dots, x_n , decompõe-se em $n-1$ graus de liberdade, cada um deles correspondendo a uma componente $\frac{X_i^2}{N_i}$ da variação entre os n números.

* Denomina-se "contraste", uma forma linear cuja soma dos coeficientes é igual a zero; assim, $3x_1 - x_2 - x_3 - x_4$ é um contraste.

⁶ A norma da forma linear $a_1 x_1 + a_2 x_2 + a_3 x_3 + \dots + a_n x_n$ é a soma dos quadrados dos coeficientes da forma, isto é, $a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2$.

QUADRO 1. Precipitações mensais, ajustadas a um mês de 30 dias (Pelotas, 1900/1951)

	1000	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
Jan.	59,3	75,0	100,6	34,6	150,9	21,6	19,5	15,0	117,5	143,7	197,0	114,0	78,7	48,5	191,3	128,3	125,9	70,8
Fev.	98,8	131,9	212,8	214,3	224,3	165,5	11,3	104,6	118,4	107,1	57,2	65,6	126,2	82,0	76,5	165,4	51,2	113,5
Mar.	181,5	52,0	176,1	36,1	69,1	1,0	48,4	271,6	28,4	98,7	118,6	39,7	66,9	146,0	276,0	122,3	66,2	93,7
Abr.	61,7	47,9	152,1	108,0	79,4	104,9	218,7	19,9	146,3	65,0	171,6	230,7	130,7	39,6	146,2	126,3	51,6	39,9
Mai.	139,5	100,5	137,5	55,7	112,5	105,1	95,8	44,0	150,2	21,5	12,8	50,6	342,7	46,4	43,6	202,4	70,5	24,4
Jun.	28,7	22,6	56,4	219,9	175,9	68,9	256,9	161,3	136,9	200,7	24,1	87,5	146,0	42,4	287,6	50,7	75,2	30,2
Jul.	165,0	61,4	281,5	54,6	70,0	200,4	47,9	73,6	95,9	160,2	63,8	28,1	133,0	144,2	313,6	135,1	171,8	41,0
Ago.	211,3	64,8	142,7	174,7	154,4	188,6	219,2	67,0	128,5	15,9	91,8	178,2	154,3	69,0	139,5	79,9	138,6	51,8
Set.	110,7	225,7	136,7	70,3	88,0	223,4	55,8	180,7	171,0	123,8	158,9	144,4	26,7	86,9	210,8	184,8	94,0	43,8
Out.	89,4	63,7	64,9	72,6	169,5	145,6	128,7	55,4	56,6	49,3	37,3	14,5	81,0	82,1	260,2	70,9	37,0	31,3
Nov.	35,3	40,7	217,8	149,6	75,8	65,0	53,1	17,6	104,6	77,9	42,1	163,6	92,2	45,3	242,4	112,5	40,3	36,2
Dez.	117,0	42,2	89,2	150,7	124,3	193,0	38,2	133,9	34,4	112,8	14,7	189,4	132,4	39,1	111,8	60,5	65,7	105,8
	1298,2	928,4	1772,3	1321,1	1484,1	1483,0	1193,5	1144,6	1288,7	1185,6	989,9	1307,3	1510,8	871,5	2299,5	1439,1	988,0	682,4
	1918	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Jan.	83,8	22,4	78,9	63,6	153,9	176,1	27,9	125,4	32,3	80,0	57,4	115,8	93,9	231,9	30,5	57,1	146,8	44,2
Fev.	255,2	53,4	196,7	105,0	227,4	68,7	47,5	267,2	5,4	62,3	120,5	148,6	118,4	55,0	242,6	182,6	232,4	183,3
Mar.	73,3	17,1	22,3	132,1	22,6	41,6	236,8	29,1	34,3	85,5	93,9	116,2	131,6	137,0	269,5	26,1	140,5	173,5
Abr.	61,0	105,9	86,1	39,8	82,0	106,2	155,2	37,8	131,1	181,6	113,9	5,8	105,2	76,8	216,7	31,0	25,5	162,5
Mai.	128,2	57,1	69,7	122,4	63,8	32,8	42,9	194,3	41,2	43,5	104,1	94,4	116,5	210,3	160,2	35,1	177,6	25,9
Jun.	109,8	120,5	328,5	127,0	65,9	37,8	62,8	0,0	145,0	121,5	184,4	76,2	137,3	166,8	57,0	61,1	162,9	162,4
Jul.	101,6	253,1	70,6	20,8	42,1	119,8	64,7	56,3	237,6	105,6	142,9	83,9	59,0	100,7	96,8	186,9	45,7	72,4
Ago.	42,1	89,0	26,6	77,8	147,8	211,8	116,1	146,8	64,1	79,4	73,8	292,6	179,5	41,8	153,9	105,2	100,0	91,9
Set.	288,6	76,0	36,0	110,5	89,2	158,0	122,9	163,0	104,9	91,7	103,6	125,8	110,2	17,5	174,0	128,0	173,3	225,2
Out.	173,4	71,0	64,4	177,9	45,9	46,0	6,5	189,6	136,2	150,1	83,3	150,7	53,9	34,4	165,9	44,4	149,0	141,9
Nov.	51,6	154,6	126,0	56,4	55,7	154,0	24,8	52,2	98,1	47,7	9,5	26,1	87,8	180,6	53,8	25,5	82,4	52,3
Dez.	79,5	100,1	92,6	5,9	116,1	185,2	7,8	113,9	51,6	100,2	41,0	41,0	119,4	22,2	82,7	85,5	84,9	148,9
	1448,1	1120,2	1197,4	1039,2	1113,4	1338,0	915,9	1375,6	1079,8	1149,1	1128,3	1276,1	1312,7	1274,8	1703,6	968,5	1521,0	1484,4
	1936	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	Soma	Média
Jan.	62,4	62,4	224,5	89,4	260,3	262,5	66,1	24,3	236,9	42,7	110,0	98,5	329,3	85,0	10,1	125,8	5403,0	103,9
Fev.	183,9	302,1	72,3	124,8	100,4	182,3	101,8	107,8	58,3	74,0	234,4	84,8	95,4	97,5	257,1	108,5	6883,5	132,5
Mar.	94,4	184,0	137,4	100,1	75,2	42,0	168,2	37,0	293,3	27,7	15,3	32,7	81,5	203,2	107,9	141,2	5513,4	106,0
Abr.	221,1	14,6	92,0	145,9	202,9	274,1	130,6	3,7	43,6	55,8	39,6	48,0	92,0	25,0	84,0	76,6	5273,1	101,4
Mai.	375,9	104,7	25,9	55,6	161,1	263,8	232,9	82,8	21,8	53,4	59,5	85,5	36,2	34,0	174,9	53,9	5297,6	101,9
Jun.	214,9	120,1	141,7	212,0	141,8	58,1	124,3	69,1	102,8	58,2	126,6	75,5	191,0	73,2	129,8	70,1	6107,8	117,5
Jul.	65,7	271,7	125,1	21,0	129,0	107,4	73,1	67,3	19,9	121,5	114,1	64,3	111,6	23,7	75,4	17,7	5519,1	108,1
Ago.	158,5	182,3	70,2	157,5	25,6	421,5	126,4	40,5	72,7	154,5	127,9	54,8	12,3	87,5	65,3	57,5	6126,4	117,8
Set.	93,2	202,0	219,6	114,5	97,4	51,2	29,3	73,8	64,6	88,5	146,6	212,2	249,8	148,2	105,6	130,7	6661,0	128,1
Out.	324,4	86,9	67,0	90,4	236,1	50,6	127,3	34,0	151,5	41,3	161,9	44,1	91,0	173,1	54,2	47,6	5175,9	99,5
Nov.	44,9	69,5	84,4	86,4	86,6	115,8	13,2	86,6	44,4	86,4	99,3	67,1	103,2	10,0	48,2	144,2	4089,3	78,6
Dez.	121,5	6,8	63,2	63,5	194,4	83,9	23,1	48,7	17,8	57,2	64,9	52,8	29,6	71,5	110,7	46,2	4269,4	82,1
	1960,8	1712,1	1292,3	1261,1	1690,8	1913,2	1216,3	675,6	1127,6	861,2	1300,1	910,3	1422,9	1097,9	1223,2	1020,0	66319,5	

APLICAÇÕES NA ANÁLISE HARMÔNICA

Exemplificar-se-á com as precipitações mensais em Pelotas, de janeiro a dezembro, nos anos de 1900 a 1951⁷, ajustadas a um mês de trinta dias, transcritas no Quadro 1.

Embora a precipitação em cada mês não seja a mesma nos anos sucessivos, pode-se presumir que as diferenças correm à conta de variações anuais que se superpõem a um movimento periódico anual.

⁷ O ajustamento foi feito multiplicando a precipitação mensal por 30/31, nos meses com 31 dias, por 30/28 no mês de fevereiro dos anos comuns e por 30/29 no mês de fevereiro dos anos bissextos, reduzindo-se, assim, tôdas as precipitações mensais a um mês de trinta dias.

Uma estimativa desse movimento periódico anual é dada na última coluna do Quadro 1, onde se acham representadas as médias das precipitações de cada mês nos 52 anos decorridos de 1900 a 1951. Não se discutirá, no momento, a validade dessa estimativa, no caso das precipitações mensais.

A análise harmônica permite decompor uma variação periódica de período T , em ondas senoidais de períodos T , $T/2$, $T/3$, etc.

No caso das precipitações mensais, tem-se $T = 12$ meses, considerando-se o mês como unidade de tempo; a variação anual pode decompor-se, neste caso, em seis ondas senoidais de períodos de

QUADRO 2. Totais das precipitações mensais (1900/51) e coeficientes para a análise harmônica

Mês	i	x_i	cos 30i	sen 30i	cos 60i	sen 60i	cos 90i	sen 90i	cos 120i	sen 120i	cos 150i	sen 150i	cos 180i
J	0	5 403,0	1	0	1	0	1	0	1	0	1	0	1
F	1	6 833,5	0,866	0,5	0,5	0,866	0	1	-0,5	0,866	-0,866	0,5	-1
M	2	5 513,4	0,5	0,866	-0,5	0,866	-1	0	-0,5	-0,866	0,5	-0,866	1
A	3	5 273,1	0	1	-1	0	0	-1	1	0	0	1	-1
M	4	5 297,6	-0,5	0,866	-0,5	-0,866	1	0	-0,5	0,866	-0,5	-0,866	1
J	5	6 107,8	-0,866	0,5	0,5	-0,866	0	1	-0,5	-0,866	0,866	0,5	-1
J	6	5 519,1	-1	0	1	0	-1	0	1	0	-1	0	1
A	7	6 126,4	-0,866	-0,5	0,5	0,866	0	-1	-0,5	0,866	0,866	-0,5	-1
S	8	6 661,0	-0,5	-0,866	-0,5	0,866	1	0	-0,5	-0,866	-0,5	0,866	1
O	9	5 175,0	0	-1	-1	0	0	1	1	0	0	-1	-1
N	10	4 039,3	0,5	-0,866	-0,5	-0,866	-1	0	-0,5	0,866	0,5	0,866	-1
D	11	4 269,4	0,866	-0,5	0,5	-0,866	0	-1	-0,5	-0,866	-0,866	-0,5	-1

12, 6, 4, $2\frac{2}{5}$ e 2 meses, respectivamente, oscilando acima e abaixo da precipitação média anual reduzida a um período de trinta dias.

Fazendo

$$x_i = p_0 + p_1 \cos ir + p_2 \cos 2ir + p_3 \cos 3ir + p_4 \cos 4ir + p_5 \cos 5ir + p_6 \cos 6ir + q_1 \sin ir + q_2 \sin 2ir + q_3 \sin 3ir + q_4 \sin 4ir + q_5 \sin 5ir,$$

onde $r = \frac{360^\circ}{12} = 30^\circ$, $n = 12$, $i = 0, 1, 2, \dots, 11$ (subtraindo 1 do número de ordem do mês), preparou-se o Quadro 2, no qual se encontram as somas das precipitações mensais em Pelotas, de 1900 a 1951, reduzidas a um mês de trinta dias, e os coeficientes de $p_1, p_2, \dots, p_6, q_1, q_2, \dots, q_5$ para cada um dos doze meses².

Em virtude da ortogonalidade, o sistema de doze equações lineares nas doze incógnitas $p_0, p_1, p_2, \dots, p_6, q_1, q_2, \dots, q_5$ é facilmente resolvido, obtendo-se

$$p_0 = \bar{x}$$

$$p_1 = \frac{1}{6} \sum_{i=0}^{11} x_i \cos ir$$

$$p_2 = \frac{1}{6} \sum_{i=0}^{11} x_i \cos 2ir$$

$$p_3 = \frac{1}{6} \sum_{i=0}^{11} x_i \cos 3ir$$

$$p_4 = \frac{1}{12} \sum_{i=0}^{11} x_i \cos 4ir$$

$$p_5 = \frac{1}{6} \sum_{i=0}^{11} x_i \cos 5ir$$

$$q_1 = \frac{1}{6} \sum_{i=0}^{11} x_i \sin ir$$

² Representam-se na tabela, os totais e não as médias, porque se pretende fazer, adiante, uma análise conjunta dos 52 anos.

$$q_1 = \frac{1}{6} \sum_{i=0}^{11} x_i \sin 2ir$$

$$q_5 = \frac{1}{6} \sum_{i=0}^{11} x_i \sin 5ir$$

Somando os quadrados dos desvios das precipitações mensais em relação à precipitação média anual, reduzida a trinta dias, obtém-se

$$\sum_{i=0}^{11} (x_i - \bar{x})^2 = 6(p_1^2 + q_1^2 + p_2^2 + q_2^2 + \dots + p_5^2 + q_5^2 + 2p_6^2),$$

decompondo-se, assim, a variação entre os meses em 11 componentes ortogonais, cada uma delas correspondendo a um grau de liberdade.

Grupando os termos em p_j e q_j , tem-se

$$p_j \cos jir + q_j \sin jir = a_j (\sin A_j \cos jir + \cos A_j \sin jir) = a_j \sin (jir + A_j)$$

onde $a_j^2 = p_j^2 + q_j^2$ e $A_j = \arctg \frac{p_j}{q_j}$,

o quadrante onde se encontra A_j sendo determinado pelos sinais do seno e do cosseno (sinais de p_j e de q_j), respectivamente.

Quanto ao termo em p_6 , tem-se

$$p_6 \cos 6ir = a_6 \sin (6ir + A_6)$$

onde $a_6 = |p_6|$ e A_6 é igual a 90° ou a 270° , conforme seja p_6 positivo ou negativo.

No caso em estudo, sendo $n = 12$, tem-se $r = 30^\circ$ e, assim

$$x_i - \bar{x} = a_1 \sin (30i + A_1) + a_2 \sin (60i + A_2) + a_3 \sin (90i + A_3) + a_4 \sin (120i + A_4) + a_5 \sin (150i + A_5) + a_6 \sin (180i + A_6),$$

sendo A_6 igual a 90° ou a 270° , conforme seja p_6 positivo ou negativo.

A variação anual da precipitação decompõe-se, dessa forma, em seis ondas senoidais, da forma

$$a_j \text{sen}(30ji + A_j)$$

de período igual a 12/j meses, de amplitude a_j e de fase A_j .

A soma dos quadrados das diferenças entre as precipitações mensais e a média da precipitação anual, reduzida a trinta dias, exprime-se em função das amplitudes das ondas senoidais, na forma

$$\sum_{i=0}^{11} (x_i - \bar{x})^2 = 6(a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2) + 12a_6^2$$

correspondendo 2 graus de liberdade a cada uma das cinco primeiras parcelas e 1 grau de liberdade à última componente, 12 a_6^2 .

MARCHA DA ANÁLISE

Dividindo por 12 a soma dos totais mensais das precipitações mensais (Quadro 2), tem-se

$$p_0 = \frac{1}{12} S x_i = 5\,526,63.$$

A precipitação média anual, reduzida a trinta dias, seria obtida dividindo p_0 pelo número de anos, 52 no caso. Preferiu-se, entretanto, trabalhar com os totais, e não com as médias, para utilizar, adiante, aqueles totais nos testes de significância das amplitudes das componentes harmônicas.

Dividindo por 6 a soma dos produtos das precipitações mensais pelos valores correspondentes de $\cos 30i$ e, respectivamente, de $\text{sen } 30i$ (Quadro 2), obtém-se

$$p_1 = \frac{1}{6} S x_i \cos(30i) = -371,74$$

$$q_1 = \frac{1}{6} S x_i \text{sen}(30i) = 241,25$$

e, de maneira análoga,

$$p_2 = \frac{1}{6} S x_i \cos(60i) = 231,00$$

$$q_2 = \frac{1}{6} S x_i \text{sen}(60i) = 782,32$$

$$p_3 = \frac{1}{6} S x_i \cos(90i) = 373,30$$

$$q_3 = \frac{1}{6} S x_i \text{sen}(90i) = 416,38$$

$$p_4 = \frac{1}{6} S x_i \cos(120i) = -183,85$$

$$q_4 = \frac{1}{6} S x_i \text{sen}(120i) = -22,34$$

$$p_5 = \frac{1}{6} S x_i \cos(150i) = -59,61$$

$$q_5 = \frac{1}{6} S x_i \text{sen}(150i) = 223,73$$

$$p_6 = \frac{1}{12} S x_i \cos(180i) = -112,73.$$

Note-se que, no cálculo de p_6 , o divisor é 12 e não 6 como nos demais.

Com esses dados foram calculados as amplitudes e os ângulos fase das componentes harmônicas:

$$a_1 = \sqrt{p_1^2 + q_1^2} = 443,16, \quad A_1 = \text{arc tg} \frac{p_1}{q_1} = 302^\circ 59'$$

$$a_2 = \sqrt{p_2^2 + q_2^2} = 815,71, \quad A_2 = \text{arc tg} \frac{p_2}{q_2} = 16^\circ 27'$$

$$a_3 = \sqrt{p_3^2 + q_3^2} = 559,22, \quad A_3 = \text{arc tg} \frac{p_3}{q_3} = 41^\circ 51'$$

$$a_4 = \sqrt{p_4^2 + q_4^2} = 185,20, \quad A_4 = \text{arc tg} \frac{p_4}{q_4} = 263^\circ 04'$$

$$a_5 = \sqrt{p_5^2 + q_5^2} = 231,54, \quad A_5 = \text{arc tg} \frac{p_5}{q_5} = 345^\circ 05'$$

$$a_6 = |p_6| = 112,73, \quad A_6 = 270^\circ$$

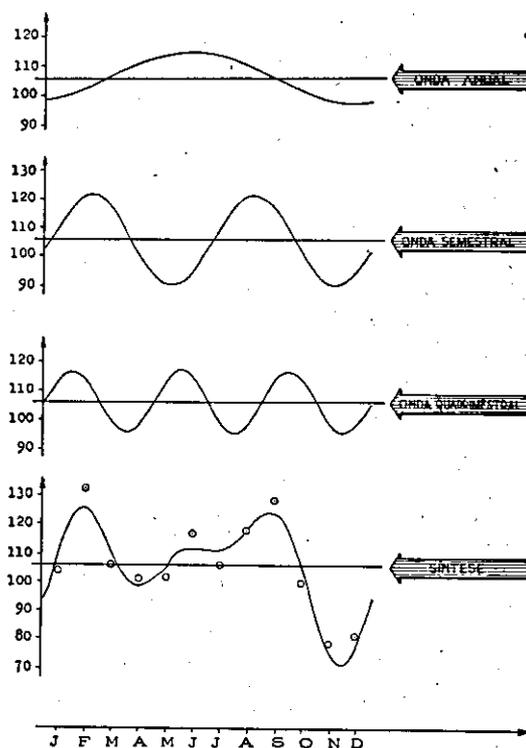


FIG. 7. Variação anual das precipitações mensais em Pelotas, 1900-1951.

Na Fig. 7, estão representadas as harmônicas de períodos de 12, 6 e 4 meses, que foram significativas, como se verá adiante, $a_j \text{sen}(30ji + A_j)$, $j = 1, 2$ e 3, e a síntese dessas ondas, $a_1 \text{sen}(30i + A_1) + a_2 \text{sen}(60i + A_2) + a_3 \text{sen}(90i + A_3)$, onde as ordenadas são os desvios em relação à média anual ($p_0/52 = 5\,526,63/52 = 106,28$) e as amplitudes foram divididas por 52, para reduzi-las a um ano médio.

TESTE DE SIGNIFICÂNCIA ⁹

Admitindo-se que as precipitações mensais se distribuam normalmente em torno das respectivas médias com a mesma variância σ^2 , demonstra-se que a distribuição das amplitudes das componentes harmônicas é dada pela equação diferencial

$$dP = \frac{Na}{2\sigma^2} e^{-Na^2/4\sigma^2} da,$$

onde a é a amplitude da onda e N o período fundamental (no caso, $N = 12$ meses)¹⁰.

Essa é a chamada "distribuição circular normal" (Brooks & Carruthers 1953), que se aplica, também na análise estatística de ventos e, em geral, de variáveis vetoriais.

A verdadeira variância, σ^2 , não é conhecida, mas pode ser estimada com base nas observações de um conjunto apreciável de anos e estendida aos anos posteriores, o que implica em admitir a estabilidade da variância.

A diferencial

$$dP = \frac{Na}{2\sigma^2} e^{-Na^2/4\sigma^2} da, \quad 0 \leq a < +\infty,$$

é, a menos de um infinitésimo de ordem superior à primeira em relação a da , a probabilidade da amplitude de uma onda casual compreendida entre a e $a + da$.

Adotando-se o nível de significância de 5% ($\alpha = 0,05$), o valor de a acima do qual a amplitude de uma onda, com dois graus de liberdade, será significativa, é dado pela equação integral

$$\int_0^a \frac{Na}{2\sigma^2} e^{-Na^2/4\sigma^2} da = 1 - \alpha = 0,95.$$

Integrando, tem-se

$$\left[-e^{-Na^2/4\sigma^2} \right]_0^a = 1 - e^{-Na^2/4\sigma^2},$$

⁹ O leitor exclusivamente interessado nos aspectos práticos do problema, poderá deixar de ler este parágrafo e os quatro seguintes, passando diretamente ao parágrafo sobre a análise da variação (pág. 28).

¹⁰ Ver Conrad e Pollak (1950), § 13.1.1.

donde

$$\alpha = e^{-Na^2/4\sigma^2} \quad \text{e} \quad a = 2\sigma \sqrt{\frac{-\text{Log } \alpha}{N}}.$$

Assim, se $\alpha = 0,05 = 1/20$,

$$a = 2\sigma \sqrt{\frac{\text{Log } 20}{N}},$$

onde $\text{Log } 20 = (\log 20)/0,434\,3$, ou

$$a = \lambda \sigma / \sqrt{N},$$

onde

$$\lambda = 2 \sqrt{\text{Log } 20}.$$

Brooks e Carruthers (1953) observam que, se diversas ondas forem calculadas, o nível de significância de 5% não é suficiente, pois, por definição, uma em vinte das amplitudes calculadas em uma sucessão excede o valor correspondente àquele nível. Nestas condições, os mesmos autores recomendam um teste de amplitude múltipla, fazendo $\lambda_n = 2 \sqrt{\text{Log } 20n}$, onde n é o número de ordem da amplitude na sucessão das amplitudes a testar, dispostas em ordem crescente.

Se uma das amplitudes não fôr significativa, também serão consideradas não significativas as ondas de menor amplitude.

Os valores de λ_n , para n de 1 a 33 e para $n = 50$, $n = 100$ e $n = 300$, encontram-se no Quadro 3.

O estabelecimento da fórmula $\lambda_n = 2 \sqrt{\text{Log } 20n}$ encontra-se no Apêndice, item 8, no presente trabalho.

O teste de amplitude múltipla acima recomendado obedece ao critério de Newman e Keuls, citado por Cochran e Cox (1960), § 3.53.

Como observa Duncan (1955), quando testes múltiplos são independentes entre si, como ocorre no caso que está sendo considerado, não há por que recorrer a um teste de amplitude múltipla: o uso conjunto de testes independentes, com o mesmo nível de significância, é tão válido como se os dados que servem de base a cada um deles proviessem de observações distintas. Segundo Duncan (1955), portanto, tôdas as amplitudes em uma mesma análise harmônica seriam testadas comparando cada uma delas com $a = \lambda \sigma / \sqrt{N}$, onde $\lambda = \lambda_1 = 2 \sqrt{\text{Log } 20} = 3,46$.

QUADRO 3. Teste de significância de amplitudes de ondas harmônicas

n	1	2	3	4	5	6	7	8	9	10
λ_n	3,46	3,84	4,05	4,19	4,29	4,38	4,45	4,51	4,56	4,60
n	11	12	13	14	15	16	17	18	19	20
λ_n	4,64	4,68	4,72	4,75	4,78	4,80	4,83	4,85	4,88	4,90
n	21	22	23	24	25	26	27	28	29	30
λ_n	4,92	4,93	4,95	4,97	4,99	5,00	5,02	5,03	5,04	5,06
n	31	32	33	34	35	36	...	50	100	300
λ_n	5,07	5,08	5,10	5,11	5,12	5,13	...	5,26	5,51	5,90

O QUADRADO DA AMPLITUDE

Alguns autores apresentam periodogramas onde as abscissas são os períodos das ondas harmônicas ($N, N/2, N/3, \dots$) e as ordenadas os quadrados das amplitudes respectivas. Fora desse caso, onde o quadrado das amplitudes é usado para tornar mais nítidas as diferenças entre elas, trabalha-se, geralmente, com a própria amplitude.

Parece ao autor, entretanto, como será adiante justificado, que há toda a conveniência em operar com o quadrado da amplitude.

Substituindo em

$$dP = \frac{Na}{2\sigma^2} e^{-Na^2/4\sigma^2} da, \quad 0 \leq a < +\infty,$$

$da = da^2/2a$, tem-se a distribuição do quadrado da amplitude:

$$dP = \frac{N}{4\sigma^2} e^{-Na^2/4\sigma^2} da^2 \quad 0 \leq a^2 < +\infty,$$

que é a distribuição exponencial negativa¹¹.

Uma amplitude será significativa se exceder $2\sigma\sqrt{\text{Log } 20n/N}$, como no parágrafo precedente, ou, o que é o mesmo, se seu quadrado exceder $\lambda_n^2 \sigma^2/N$, sendo λ_n dado no Quadro 3, segundo o critério de Newmann e Keuls, citado por Cochran e Cox (1960), ou $\lambda^2 \sigma^2/N = \lambda_1^2 \sigma^2/N$, segundo o critério de Duncan (1955).

DISTRIBUIÇÃO DA DIFERENÇA ENTRE OS QUADRADOS DE DUAS AMPLITUDES

Uma das vantagens de operar com os quadrados das amplitudes é que o módulo da diferença entre dois deles, quando as ondas respectivas são harmônicas e as variâncias são homogêneas, tem a mesma distribuição exponencial negativa do quadrado de uma amplitude.¹² Em outras palavras, a raiz quadrada do módulo da diferença entre os quadrados das amplitudes de duas ondas harmônicas homogêneas tem a mesma distribuição de uma amplitude¹².

Nestas condições, pode-se testar a significância, não só de cada um dos quadrados das amplitudes das ondas resultantes de uma análise harmônica, mas também, da diferença entre os quadrados das amplitudes de duas delas, sendo a diferença considerada significativa se exceder o quadrado de $a = \lambda_n \sigma \sqrt{1/N}$, onde n é a diferença entre os números de ordem das duas ondas, dispostas na ordem crescente das

respectivas amplitudes e λ_n é dado no Quadro 3 (critério de Newmann e Keuls, citado por Cochran e Cox 1960).

Se for considerado conveniente adotar o critério de Duncan (1955), a fim de aumentar a proteção contra erros de segunda espécie, adotar-se-á, como antes, $\lambda_n = \lambda_1 = 3,46$. Evidentemente, se a maior de duas amplitudes não for significativa, também não o será a da menor delas nem a diferença entre seus quadrados.

A RAZÃO ENTRE OS QUADRADOS DE DUAS AMPLITUDES

Os testes referidos nos parágrafos precedentes exigem o prévio conhecimento da variância dos dados originais. Quando se dispõe de uma estimativa da variância baseada em número relativamente considerável de observações, essa estimativa pode ser usada em lugar do valor desconhecido nos testes de significância. Não é esse, entretanto, um processo rigoroso, pois não toma em consideração o erro de amostragem da variância estimada.

A partir da distribuição do quadrado a^2 de uma amplitude,

$$\frac{N}{4\sigma^2} e^{-Na^2/4\sigma^2} da^2,$$

pode-se demonstrar (Apêndice, item 10) que a razão dos quadrados de duas amplitudes,

$$R = a_2^2/a_1^2,$$

se distribui como F com 2 e 2 graus de liberdade:

$$dP = \frac{1}{(1+R)^2} dR, \quad 0 \leq R < +\infty.$$

Assim, com as precipitações mensais em uma localidade em dois anos, pode-se fazer a análise harmônica em cada um dos anos e no conjunto dos dois, obtendo-se as amplitudes a'_1 , a''_1 e a_1 das ondas anuais e, semelhantemente, as das outras harmônicas. A significância da onda anual poderia ser testada comparando-se $R = \frac{a_1^2}{|a_1'^2 - a_1''^2|}$ com o valor de F significativo no nível de 5% (F-teste unilateral), com 2 e 2 graus de liberdade.

TESTE g, DE FISHER

Em 1929, Fisher (1950b) divulgou testes exatos de significância de

$$g = \frac{a_i^2}{a_1^2 + a_2^2 + \dots + a_r^2}$$

¹¹ Daí resulta que $Na^2/4\sigma^2 = \frac{1}{2}$ se distribui como χ^2 com 2 graus de liberdade.

¹² Ver Apêndice, item 9.

onde $r = I \left(\frac{N-1}{2} \right)$ é o maior interior que não ultrapassa $(N-1)/2$.

Se N é ímpar, g representa a fração da variação entre os dados originais que corre à conta da harmônica de ordem i ,

$$g = \frac{Na_i^2/2}{S(x_i - \bar{x})^2}$$

e, se N é par, a harmônica de maior frequência não é tomada em consideração, sendo

$$\frac{N}{2} (a_1^2 + a_2^2 + \dots + a_r^2) = \frac{S(x_i - \bar{x})^2 - \frac{(x_1 - x_2 + x_3 - \dots - x_N)^2}{N}}{N}$$

onde $r = N/2 - 1$.

Como mostra Fisher (1950b) a probabilidade de g , correspondente a uma harmônica escolhida *a priori*, atingir ou superar um dado valor g é

$$P = (1-g)^{n-1}$$

Se a componente harmônica não fôr escolhida *a priori*, mas submetida ao teste por ter a maior amplitude, a probabilidade correspondente é

$$P = n(1-g)^{n-1} - \frac{n!}{2!(n-2)!} (1-2g)^{n-1} + \dots + (-1)^{k-1} \frac{n!}{k!(n-k)!} (1-kg)^{n-1}$$

onde k é o maior inteiro que não supera $1/g$.

Os valores significativos de g , dados por esta última fórmula, correspondentes aos níveis de significância $\alpha = 0,05$ e $\alpha = 0,01$, encontram-se tabelados no citado trabalho.

Em novo trabalho, Fisher (1950a), apresentou o teste correspondente à segunda harmônica, na ordem das amplitudes decrescentes, a ser usado quando a significância da harmônica de maior amplitude é duvidosa e se suspeita que as duas ondas mais importantes sejam devidas a causas sistemáticas.

Os testes referidos no presente parágrafo podem ser úteis quando se dispõe de dados relativos a um só ano ou à média de uma sucessão de anos, sem se ter acesso aos dados anuais.

ANÁLISE DA VARIAÇÃO

Com as precipitações mensais em um conjunto de n anos, ter-se-ia n análises harmônicas, uma para cada ano, e uma análise no conjunto dos anos. Os contrastes entre os quadrados das amplitudes da onda anual, por exemplo, nos diversos anos, devem ser combinados de modo a fornecer uma soma de quadrados das amplitudes de ondas casuais de igual período. A possibilidade dessa combinação resulta

dos fatos, já assinalados, que $Na^2/2\sigma^2$ se distribui como χ^2 com 2 graus de liberdade, que o módulo da diferença entre os quadrados de duas amplitudes se distribui da mesma forma que o quadrado de uma amplitude e , ainda, da atividade do χ^2 , isto é, que a soma de dois qui-quadrados, com m e n graus de liberdade, se distribui como um qui-quadrado com $m+n$ graus de liberdade.

Na prática, o processo acima esboçado é equivalente a uma análise da variação de cada uma das ondas de períodos $N = 12, N/2, N/3$, etc.

Indicar-se-á a marcha da análise da onda anual com os dados de Pelotas (1900/51).

No Quadro 4, encontram-se para cada ano e para as precipitações mensais totais de r anos, os valores de

$$6p_i = \sum_{i=0}^r x_i \cos 30i \quad \text{e de} \quad 6q_i = \sum_{i=0}^r x_i \sin 30i$$

A soma dos valores de $6p_1$ nos 52 anos deve coincidir com o contraste $6p_1$ calculado no conjunto dos anos e o mesmo vale para $6q_1$.

QUADRO 4. Valores de $6p_1$ e de $6q_1$ (Pelotas, 1900/51)

ANO	$6p_1$	$6q_1$	ANO	$6p_1$	$6q_1$
1900	-143,357 2	+ 23,450 0	1926	- 344,868 6	-- 96,437 0
01	- 23,067 8	-- 60,687 4	27	-- 59,854 4	+ 24,503 6
02	- 20,438 6	+ 66,100 6	28	- 221,302 2	+ 199,173 4
03	- 33,103 6	-- 11,134 6	29	- 164,835 2	-- 149,065 8
04	+ 58,947 8	-- 13,935 2	1930	-- 37,164 0	+ 73,059 6
05	- 222,584 0	-- 272,171 8	31	+ 62,480 8	+ 250,407 2
06	- 422,885 6	+ 125,969 8	32	+ 27,320 4	+ 237,145 4
07	- 17,516 8	+ 98,581 8	33	-- 97,391 2	-- 66,831 8
08	- 170,011 6	+ 51,808 0	34	+ 84,210 4	+ 35,738 4
09	- 6,992 2	+ 34,671 0	35	+ 26,611 4	+ 5,415 4
1910	+ 89,596 0	+ 61,426 4	36	- 227,038 0	+ 243,785 2
11	+ 81,649 8	-- 79,078 2	37	- 182,771 0	+ 150,075 2
12	- 195,562 2	+ 294,196 2	38	+ 6,887 6	-- 29,700 2
13	- 58,299 8	+ 17,783 2	39	-- 80,319 2	+ 74,256 8
14	- 197,100 8	-- 173,207 6	1940	+ 183,278 4	+ 40,511 8
15	- 0,470 2	+ 116,978 4	41	- 103,304 4	+ 211,200 8
16	- 158,815 4	-- 22,271 6	42	- 156,342 8	+ 352,147 6
17	+ 179,511 6	+ 34,794 6	43	-- 18,884 6	-- 21,603 6
18	- 5,445 2	-- 110,814 2	44	+ 256,569 6	+ 105,882 6
19	- 259,896 0	- 103,142 4	45	- 163,279 0	- 106,480 8
1920	- 27,382 8	+ 163,030 0	46	-- 11,053 2	-- 186,372 6
21	- 60,717 4	+ 11,911 6	47	- 63,428 2	-- 117,922 6
22	+ 187,356 8	+ 1,005 0	48	+ 99,242 2	-- 50,519 8
23	+ 62,423 8	-- 290,811 6	49	+ 86,987 8	-- 8,640 0
24	- 95,937 6	+ 256,212 0	1950	+ 22,058 2	+ 246,964 0
25	+ 134,003 8	-- 141,448 8	51	+ 181,968 6	-- 2,656 8
				Soma	- 2 230,455 8 + 1 447,516 2

Somam-se os quadrados dos $6p_1$ nos diversos anos, obtendo-se a "variação total"; divide-se por 52 (número de anos) o quadrado do contraste $6p_1$ calculado no conjunto dos anos, obtendo-se $36p_1^2$; a variação entre os valores dos contrastes nos diversos anos é obtida por diferença. Dividindo-se as variações em cada uma das primeiras filas pelos respec-

tivos graus de liberdade (1 e 51), obtêm-se os quadrados médios.

	G. L.	S. Q.	Q. M.	F
Contraste ($36p_1^2/52$)	1	95 671,789 9	95 671,789 9	4,83*
Varição do contraste entre anos	51	1 010 356,602 5	19 810,913 7	
Varição total	52	1 106 028,392 4		

Procede-se semelhantemente com o contraste $6q_1$, obtendo-se a seguinte análise da variação:

	G. L.	S. Q.	Q. M.	F
Contraste ($36q_1^2/52$)	1	40 294,291 3	40 294,291 3	1,96—
Varição entre anos	51	1 049 804,361 0	20 584,399 2	
Varição total	52	1 090 098,652 3		

A razão de duas variâncias homogêneas, cada uma delas com 51 graus de liberdade, pode alcançar, com probabilidade de 5% valores menores do que 0,565, como se pode verificar consultando uma tabela de F bilateral (Gomes 1963). As variâncias residuais das duas componentes da onda anual são, portanto, homogêneas, sendo permitido fundir as duas análises em uma só:

	G. L.	S. Q.	Q. M.	F
Amplitude ($36a_2^2/52$)	2	135 966,081 2	67 983,040 6	3,37*
Varição da amplitude	102	2 060 160,963 5	20 197,656 5	
Varição total	104	2 196 127,044 7		

A onda anual é significativa, o que parece óbvio. A amplitude foi igual a $\sqrt{52} \times 135 966,081 2/36 = 443,16$ mm, no total de 52 anos, ou 8,52 mm, na média de 52 anos.

As análises correspondentes às demais componentes harmônicas são resumidas a seguir.

Onda semestral

	G. L.	S. Q.	Q. M.	F
Contraste ($36p_2^2/52$)	1	36 942,230 8	36 942,230 8	1,87—
Varição entre anos	51	1 009 172,819 2	19 787,702 3	
Varição total	52	1 046 115,050 0		

	G. L.	S. Q.	Q. M.	F
Contraste ($36q_2^2/52$)	1	423 704,487 9	423 704,487 9	13,10**
Varição entre anos	51	1 649 264,400 8	32 338,517 7	
Varição total	52	2 072 968,888 7		

As variâncias residuais mostraram-se homogêneas.

	G. L.	S. Q.	Q. M.	F
Amplitude ($36a_3^2/52$)	2	460 646,718 7	230 323,359 3	8,84**
Varição da amplitude	102	2 658 437,220 0	26 063,110 0	
Varição total	104	3 119 083,938 7		

A onde semestral foi significativa. A amplitude é $a_2 = 815,71$.

Onda quadrimestral

	G. L.	S. Q.	Q. M.	F
Contraste ($36r_1^2/52$)	1	96 475,077 7	96 475,077 7	3,59—
Varição entre anos	51	1 370 875,802 3	26 870,917 7	
Varição total	52	1 467 350,880 0		

	G. L.	S. Q.	Q. M.	F
Contraste ($36q_3^2/52$)	1	120 028,901 7	120 028,901 7	3,73—
Varição entre anos	51	1 639 317,608 3	32 143,484 0	
Varição total	52	1 759 346,510 0		

As variâncias residuais são homogêneas.

	G. L.	S. Q.	Q. M.	F
Amplitude ($36a_4^2/52$)	2	216 503,979 4	108 251,989 7	3,67*
Varição da amplitude	102	3 010 193,410 6	29 511,700 8	
Varição total	104	3 226 697,360 0		

A onda quadrimestral foi significativa. A amplitude é $a_3 = 559,22$.

Onda trimestral

	G. L.	S. Q.	Q. M.	F
Contraste ($36p_3^2/52$)	1	23 400,569 4	23 400,569 4	1,20—
Varição entre anos	51	998 561,660 6	19 579,640 4	
Varição total	52	1 021 962,230 0		

	G. L.	S. Q.	Q. M.	F
Contraste ($36q_3^2/52$)	1	345,600 5	345,600 5	0,01—
Varição entre anos	51	1 757 892,664 9	34 468,483 8	
Varição total	52	1 758 238,265 4		

O F obtido, menor que 1, não foi significativo, como se pode ver numa tabela de F bilateral (1 e 51 G. L.).

As variâncias residuais das duas componentes da onda trimestral são homogêneas.

	G. L.	S. Q.	Q. M.	F
Amplitude ($36a_4^2/52$)	2	23 748,169 9	11 873,084 9	0,44—
Varição da amplitude	102	2 756 454,325 5	27 024,062 0	
Varição total	104	2 780 200,495 4		

A onda trimestral não foi significativa. O F obtido não foi significativamente menor que 1 (F bilateral, 2 e 102 G. L.). O baixo valor de F deve-se, sem dúvida, à grande variação de fase, de ano para ano, da onda trimestral.

Onda de 72 dias (quinta harmônica)

	G. L.	S. Q.	Q. M.	F
Contraste ($36p_4^2/52$)	1	2 459,795 6	2 459,795 6	0,08—
Varição entre anos	51	1 521 199,452 3	29 827,440 2	
Varição total	52	1 523 659,247 9		

	G. L.	S. Q.	Q. M.	F
Contraste ($36a_3^2/52$)	1	34 653,735 9	34 653,735 9	1,44--
Varição entre anos	51	1 225 883,935 2	24 036,940 0	
Varição total	52	1 260 537,671 1		

As variâncias residuais foram homogêneas.

	G. L.	S. Q.	Q. M.	F
Amplitude ($36a_3^2/52$)	2	37 113,531 5	18 556,765 8	0,89--
Varição da amplitude	102	2 747 093,387 5	26 932,100 1	
Varição total	104	2 784 196,919 0		

A quinta harmônica não foi significativa. O F obtido foi menor que 1, embora não tenha sido significativamente menor que 1.

Onda bimestral

	G. L.	S. Q.	Q. M.	F
Contraste ($144a_6^2/52$)	1	35 188,409 4	35 188,403 4	0,62--
Varição entre anos	51	2 880 750,920 6	56 485,312 2	
Varição total	52	2 915 939,330 0		

Note-se que a onda bimestral tem apenas uma componente (1 grau de liberdade).

A sexta harmônica não foi significativa. O F obtido foi menor que 1, embora de maneira não significativa.

TESTE DE BRUNT

Como visto anteriormente, a variação total entre os meses é

$$\sum_{i=0}^{11} (x_i - \bar{x})^2 = 6(a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2) + 12a_6^2,$$

decompondo-se, assim, os 11 G.L. da variação total em dois G.L. para cada uma das cinco primeiras harmônicas e 1 G. L. para a sexta onda.

Dividindo por 12, Brunt obtém

$$\sigma_T^2 = \frac{1}{2}(a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2) + a_6^2,$$

onde σ_T^2 é a variância da amostra.

Suponha-se que foram determinadas as amplitudes da 1.^a e da 2.^a harmônicas. Ter-se-á

$$\sigma_T^2 = \frac{1}{2}(a_1^2 + a_2^2) + \sigma_0^2,$$

sendo σ_0^2 uma variância residual.

Brunt compara a variância residual

$$\sigma_0^2 = \sigma_T^2 - \frac{1}{2}(a_1^2 + a_2^2),$$

com a variância total, σ_T^2 , deixando de calcular as demais ondas se a variância residual é muito pequena em relação à total.

Na realidade, a variância devida às duas primeiras ondas é $3(a_1^2 + a_2^2)/2$, com 4 G. L. e a variância residual é $\sigma_T^2 - \frac{6}{11}(a_1^2 + a_2^2)$, com 7 G. L., onde

$$\sigma_T^2 = \frac{\sum_{i=0}^{11} (x_i - \bar{x})^2}{11}.$$

Pode-se dizer, de maneira mais correta, que o teste de Brunt compara a variação residual,

$$S(x, -\bar{x})^2 - \frac{N}{2}(a_1^2 + a_2^2)$$

com a variação total, $S(x, -\bar{x})^2$.

Deve-se notar, entretanto, que este não é um teste de significância, mas sim um teste de completicidade.

Assim, a variação entre os meses no conjunto dos anos, reduzida a um ano, $S(x - \bar{x})^2/52 = 148\,599,591\,5$ pode decompor-se do seguinte modo:

Três primeiras ondas	6	135 519,463 2	22 586,577 2	8,63*
Desvio	5	13 080,128 3	2 616,025 7	
Varição entre os meses	11	148 599,591 5		

A variância residual, correspondente às três últimas ondas, foi, portanto, significativamente menor que a correspondente às três primeiras ondas, dispensando a determinação das ondas de período menor de quatro meses.

A variação residual (13 080,128 3) foi, neste caso, 8,80% da variação total (148 599,591 5).

A vantagem do teste de completicidade consiste em dispensar a determinação das amplitudes de tôdas as ondas, sempre que as ondas de maiores períodos sejam as mais relevantes, como neste caso.

Determinando-se tôdas as ondas, como foi feito, pode-se recorrer a um teste de significância do conjunto das ondas que não se revelaram significativas, comparando-se sua variação com a variância global média, entre anos, dessas ondas.

No caso presente, as variâncias correspondentes aos cinco graus de liberdade residuais (19 579, 640 4; 34 468, 483 6; 29 827, 440 2; 24 036, 940 0; 28 242, 656 1)¹³ foram homogêneas (teste de Bartlett, Apêndice, item 11).

A variação sistemática e a variação entre anos das três últimas ondas encontram-se a seguir.

¹³ Note-se que $12a_6^2$ e $36a_m^2$, $m < 6$, são estimativas de σ_0^2 , como se pode comprovar, recordando que

$$S(x - \bar{x})^2 = 6(a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2) + 12a_6^2.$$

Resulta que $144a_6^2/52$ e $36a_m^2/(2 \times 52)$ são estimativas de $12\sigma_0^2/52$ e de $6\sigma_0^2/52$, respectivamente, devendo a primeira ser dividida por 2 para ser comparável à segunda.

	G. L.	S. Q.
Onda trimestral	2	23 746,160 9
Onda de 12/5 meses	2	37 113,531 5
Onda bimestral	1	17 594,204 7 ¹⁴
	5	78 453,906 1

Varição residual	G. L.	S. Q.	O. M.
Onda trimestral	102	2 756 454,325 5	
Onda de 12/5 meses	102	2 747 093,387 4	
Onda bimestral	51	1 440 375,460 3 ¹⁴	
	255	6 943 913,173 1	27 231,032 0

Atribuindo tôda a variação correspondente às três últimas ondas a um só grau de liberdade, verifica-se que a variação residual¹⁵ não foi significativa.

Realmente, tem-se $F = \frac{78\ 453,906\ 1}{27\ 231,032\ 0} = 2,88$, ao passo que, para 1 e ∞ graus de liberdade, o mínimo significativo de F seria 3,84, no nível de probabilidade de 5%.

As três primeiras ondas reproduzem, pois, de maneira satisfatória, a variação anual das precipitações mensais em Pelotas, em relação à instabilidade do fenômeno nos diversos anos.

SINTESE DAS ONDAS SIGNIFICATIVAS

As somas algébricas

$$x = r_0 + a_1 \text{sen}(30i + A_1) + a_2 \text{sen}(60i + A_2) + a_3 \text{sen}(90i + A_3), \quad 0 \leq i < 12,$$

permitem traçar a curva na Fig. 7, que dá o curso esperado da variação anual das precipitações mensais em Pelotas.

Os pequenos círculos na mesma figura representam os valores observados (médias de 52 anos).

TESTE DE NORMALIDADE

É sabido que as precipitações mensais e, de modo mais geral, as precipitações em períodos curtos, não se distribuem normalmente. Nestas condições, as médias não são representativas, exigindo-se uma prévia transformação dos dados que normalize a distribuição e homogeneize as variâncias.

É possível, entretanto, que embora os dados originais não se distribuam normalmente, um ou mais contrastes entre os dados tendam à distribuição normal. Não se pode presumir, *a priori*, que os 11 contrastes considerados na análise harmônica das preci-

pitagens mensais tenham todos a mesma variância, mesmo que todos eles se distribuam normalmente. Daí o cuidado que se teve de só fazer análises conjuntas da variância depois de testar a homogeneidade das variâncias.

Quando se dispõe, como no presente caso, de dados de um número suficientemente grande de anos (52), o melhor teste de normalidade consiste em comparar as estimativas

$$g_1 = k_2/k_1^{3/2} \text{ e } g_2 = k_4/k_3^2, \text{ onde } k_2 = \frac{S(x - \bar{x})^2}{n - 1},$$

$$k_3 = \frac{nS(x - \bar{x})^3}{(n - 1)(n - 2)} \text{ e } k_4 = \frac{n}{(n - 1)(n - 2)(n - 3)}$$

$$\left[(n + 1)S(x - \bar{x})^4 - \frac{3(n - 1)}{n}S^2(x - \bar{x})^2 \right],$$

com seus respectivos desvios padrão.

$$\sqrt{\frac{6n(n - 1)}{(n - 2)(n + 1)(n + 3)}} \text{ e } \sqrt{\frac{24n(n - 1)^2}{(n - 3)(n - 2)(n + 3)(n + 5)}}$$

As estimativas g_1 e g_2 constituem medidas de assimetria e de curtose, respectivamente.

Em grandes amostras, g_1 e g_2 distribuem-se normalmente com médias iguais a zero e desvios-padrão dados pelas fórmulas acima.

Os detalhes do método podem ser encontrados em Fisher (1941).

À guisa de exemplo, aplica-se, a seguir, o teste de normalidade acima referido aos valores de

$$6p_i = \sum_{i=0}^{11} x_i \cos 30i$$

de um dos contrastes da onda anual das precipitações mensais em Pelotas nos 52 anos (1900/51). Os 52 valores, com erro menor que 0,5, foram somados a 500 a fim de evitar números negativos; os resultados obtidos foram os seguintes:

357	442	155	420
472	303	440	683
480	500	279	302
467	341	335	314
559	320	463	481
277	495	562	757
77	240	527	337
482	473	403	489
330	439	584	437
493	687	527	599
590	562	273	587
582	404	317	522
304	634	507	682

com estes dados, foram calculados:

$$S_2 = S(x - \bar{x})^2 = Sx^2 - \frac{1}{52}S^2x = 977\ 315$$

¹⁴ Dividida por 2 como anteriormente explicado.

¹⁵ Por variação residual entende-se, neste trecho, a parte correspondente às três últimas ondas.

$$S_3 = S(x - \bar{x})^3 = Sx^3 - \frac{3}{52} Sx^2 Sx + \frac{2}{52^2} Sx S^2 = -29\,484\,332$$

$$S_4 = S(x - \bar{x})^4 = Sx^4 - \frac{4}{52} Sx^3 Sx + \frac{6}{52^2} Sx^2 S^2 x = 55\,529\,922\,475$$

$$k_2 = \frac{1}{51} S_2 = 19\,163,039$$

$$k_3 = \frac{52}{51 \times 50} S_3 = -601\,219,12$$

$$k_4 = \frac{52}{51 \times 50 \times 49} \left(53S_4 - \frac{3 \times 51}{52} S_2^2 \right) = 55\,248\,823,569$$

$$g_1^2 = k_3^2 / k_2^3 = 0,051\,371$$

$$g_1 = -0,227, \quad g_2 = k_4 / k_2^2 = 0,150,$$

$$\sigma_{e1} = 0,330, \quad \sigma_{e2} = 0,650.$$

Como g_1 e g_2 foram menores, em valor absoluto, que seus respectivos desvios-padrão, conclui-se que a distribuição de p_1 não se desviou significativamente da normalidade.

Os cálculos acima resumidos são bastante trabalhosos, convindo, por isso, aplicar o teste de desvio da normalidade à distribuição de frequência dos dados grupados.

A melhor regra para determinar o número de classes parece ser a seguinte: adotar o maior número de intervalos de classe compatível com a condição de que só haja classes vazias nos extremos da distribuição. A regra refere-se à construção de distribuições de frequências grupadas para fins de teste de significância. Para representação gráfica é preferível recorrer à regra de Sturges, (citado por Waugh § 3.16): $m = 1 + 3,3 \log n$, onde m é o número de classes e n os dados a distribuir. Assim, para $n = 52$, temos $m = 1 + 3,3 \times 1,7 \approx 7$.

Na primeira coluna, acham-se os extremos dos intervalos de classe, semi-fechados à esquerda; na segunda coluna, as frequências correspondentes. Tomando como origem o ponto meio do intervalo $438 | \text{---} 484$ e como unidade a amplitude dos intervalos de classe (46), encontram-se na terceira coluna os produtos das frequências pelos desvios em relação à origem. Cada uma das colunas seguintes é obtida da precedente, multiplicando os números que nesta figuram pelos respectivos desvios em relação à origem.

Os ajustamentos que aparecem na terceira fila da parte inferior do Quadro 5 são as correções de Sheppard; $-1/12$ para k_2 e $1/120$ para k_3 .

QUADRO 5. Distribuição de frequência dos dados analisados, com 15 intervalos de classe e com amplitude igual a 96

Intervalo	Frequência				
70 --- 116	1	— 8	64	— 512	4 096
116 --- 162	1	— 7	49	— 343	2 401
162 --- 208	0	0	0	0	0
208 --- 254	1	— 5	25	— 125	625
254 --- 300	3	— 12	48	— 192	768
300 --- 346	9	— 27	81	— 243	729
346 --- 392	1	— 2	4	— 8	16
392 --- 438	5	— 5	5	— 5	5
438 --- 484	10	0	0	0	0
484 --- 530	8	8	8	8	8
530 --- 576	3	6	12	24	48
576 --- 622	5	15	45	135	405
622 --- 668	1	4	16	64	256
668 --- 714	3	15	75	375	1 875
714 --- 760	1	6	36	216	1 296
Σ	52	— 12	468	— 606	12 528
k			— 2,8	324,0	— 559,4
Ajust.				— 1,3	149,8
k^1					— 0,4
g					
S		465,2	— 283,3		12 117,7
k		9,122	— 5,777		2,285
Ajust.		— 0,033			0,008
k^1		9,039	— 5,777		2,293
g			— 0,213		0,028

Os desvios-padrão de g_1 e g_2 são, como antes, $\sigma_{e1} = 0,330$ e $\sigma_{e2} = 0,650$. Como $g_1 = -0,213$ e $g_2 = 0,028$, o desvio da normalidade não foi significativo.

Os valores de g_1 e g_2 agora obtidos não concordam muito bem com os anteriormente encontrados ($g_1 = -0,227$ e $g_2 = 0,150$). A menos que tenha havido erro de cálculo, a comparação dessas estimativas dá uma idéia da aproximação que pode dar o grupamento.

TESTE GRÁFICO DE NORMALIDADE — ORDENITOS (RANKITS)

Quando o número de dados é relativamente pequeno, não mais se pode garantir a normalidade da distribuição de g_1 e de g_2 .

Como explicado por Bliss (1958), a normalidade dos dados pode ser testada graficamente, representando em abscissas os próprios dados e, em ordenadas, os ordenitos (*rankits*)¹⁶.

Em Fisher e Yates (1949), tabela XX, para n de 2 a 50, encontram-se os ordenitos de ordem $m = 1,2,3, \dots$

À guisa de ilustração, aplica-se o método para testar a normalidade da distribuição de $6p1$, correspon-

¹⁶ Ordenito de ordem m é o desvio médio do termo de ordem m da sucessão de n valores normalmente distribuídos com média zero e desvio-padrão unidade.

dente à onda anual das precipitações mensais em Pelotas (1900/51), embora não seja êste o melhor teste com tão grande número de dados. Para isso, foi extrapolada a tabela de Fisher e Yates (1949), construindo-se as colunas correspondentes a $n = 51$ e $n = 52$:

N.º de ordem (m)	Tamanho da amostra (n)			
	49	50	51	52
1	2,24	2,25	2,26	2,27
2	1,85	1,86	1,87	1,88
3	1,62	1,63	1,64	1,65
4	1,45	1,46	1,47	1,48
5	1,32	1,33	1,35	1,36
6	1,21	1,22	1,23	1,24
7	1,11	1,12	1,13	1,15
8	1,02	1,03	1,04	1,06
9	0,94	0,95	0,97	0,98
10	0,86	0,87	0,89	0,90
11	0,79	0,80	0,82	0,84
12	0,72	0,74	0,76	0,78
13	0,66	0,67	0,69	0,71
14	0,59	0,61	0,63	0,63
15	0,53	0,55	0,57	0,59
16	0,48	0,49	0,51	0,53
17	0,42	0,44	0,46	0,48
18	0,36	0,38	0,40	0,43
19	0,31	0,33	0,36	0,38
20	0,26	0,28	0,31	0,33
21	0,21	0,23	0,26	0,28
22	0,15	0,18	0,21	0,23
23	0,10	0,13	0,16	0,18
24	0,05	0,08	0,11	0,13
25	0	0,03	0,06	0,08
26	-0,05	-0,03	0	0,03

Os pontos obtidos dispõem-se, aproximadamente, ao longo de uma reta, como se pode ver na Fig. 8, sugerindo a normalidade da distribuição.

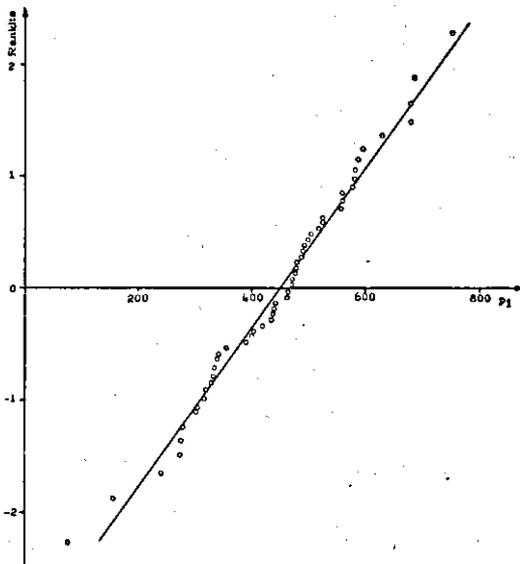


FIG. 8. Teste gráfico de normalidade.

TRANSFORMAÇÃO PRÉVIA DOS DADOS A ANALISAR

Quando os dados não se distribuem normalmente, é, às vèzes, possível obter uma transformação dos dados que normalize a distribuição.

Uma das propriedades características da distribuição normal é que as estimativas da média e da variância se distribuem independentemente.

Se fôr possível exprimir a variância, ou o desvio-padrão, em função da média, não é difícil, em geral, determinar uma transformação dos dados que homogeneíze as variâncias, isto é, que torne a variância independente da média. Admitindo-se que com essa transformação se normalize a distribuição, os dados transformados são submetidos à análise estatística adequada a dados normalmente distribuídos.

Assim, por exemplo, se a variância de x fôr proporcional à média, a variância de $z = \sqrt{x}$ é praticamente independente de x e, portanto, de z . Se a variância de x fôr proporcional ao quadrado da média, isto é, se o desvio padrão fôr proporcional à média, a transformação adequada é $z = \log x$; se, entretanto x assumir algum valor muito pequeno, ou mesmo nulo, adota-se a transformação $z = \log(x+1)$.

Analisando dados relativos às precipitações mensais em Pelotas e Ponta Grossa, Amaral e Mota (1956) usaram a transformação $z = \log\left(x + \frac{b}{a}\right)$ onde a e b são os coeficientes da equação de regressão do desvio-padrão em relação à média,

$$\sigma_x = ax + b.$$

Ao autor não parece, hoje, que no caso das precipitações mensais, o desvio-padrão seja função linear da média, parecendo-lhe preferível exprimir o desvio-padrão na forma

$$\sigma_x = Bx^a,$$

que se ajusta melhor aos dados.

Tem-se, portanto,

$$\log \sigma_x = a \log x + \log B$$

e, fazendo

$$\log x = X, \quad \log \sigma_x = Y \quad \text{e} \quad \log B = b:$$

$$Y = aX + b,$$

determinando-se a e b por regressão linear de $Y = \log \sigma_x$ em relação a $X = \log x$ ¹⁷.

Nestas condições, pode-se demonstrar que a variância de $z = x^{1-a}$ é praticamente independente de x e, portanto, de z (Apêndice, item 11).

¹⁷ "Of course, in some lines of work, the weight of y is approximately inversely proportional to y^a , whence the weight of $\log y$ is practically constant, independent of y ". (Deming 1964), Sec. 67, Exerc. 18, in fine.

PRECIPITAÇÕES MENSAIS EM MORRO VELHO (MG)

Como exemplo, serão estudadas as precipitações mensais em Morro Velho (MG), extraídas de um trabalho de Memória (1955a).

Os dados referem-se ao período 1855/1951, a maior série contínua registrada no país, segundo o autor citado.

Foram as seguintes as médias mensais e os respectivos desvios-padrão:

	J	F	M	A	M	J	J	A	S	O	N	D	Soma	Média
x	320	230	180	66	30	12	10	18	57	134	233	350	1640	136,7
σ_x	187	132	105	44	28	17	14	25	41	61	84	135		

A análise harmônica das precipitações mensais, no conjunto dos 97 anos, deu os seguintes resultados: $p_1 = 158,22$ $p_2 = 30,83$ $p_3 = -2,67$ $p_4 = -4,17$ $p_5 = -0,55$ $p_6 = 1,67$ $q_1 = -33,38$ $q_2 = -20,21$ $q_3 = -9,67$ $q_4 = -12,70$ $q_5 = -10,29$.

Nota-se, desde logo, que o curso anual das precipitações mensais, no conjunto dos 97 anos, corre quase inteiramente à conta da onda anual, cujo quadrado da amplitude é $a_1^2 = p_1^2 + q_1^2 = 26\ 147,79$.

A contribuição da onda anual na variação das precipitações mensais é, portanto, $6a_1^2 = 156\ 887$, o que representa 93,7% da variação total entre meses, $S(x - \bar{x})^2 = 167\ 385$.

Obtidos p_1 e q_1 , foram calculados os desvios em relação à média anual das precipitações mensais correspondentes à onda anual:

J	F	M	A	M	J
158,22	120,33	50,20	-33,38	-108,02	-153,71
J	A	S	O	N	D
-158,22	-120,33	-50,20	33,38	103,02	153,71

Somando aos desvios a média mensal, 136,7, têm-se as seguintes precipitações mensais em milímetros:

J	F	M	A	M	J
295	257	187	103	20	-17
J	A	S	O	N	D
-21	16	87	170	245	295

Como se vê, resultaram precipitações negativas em junho e julho, o que é absurdo e decorre, sem dúvida, da não normalidade da distribuição.

Deve-se, pois, fazer uma prévia transformação dos dados, antes de submetê-los à análise harmônica.

De $\sigma_x = Bx^a$

tem-se, como visto,

$Y = aX + b$,

sendo

$X = \log x$, $Y = \log \sigma_x$ e $b = \log B$

Calcula-se:

	x	σ_x	X	Y
J	320	187	2,505	2,272
F	230	132	2,362	2,121
M	180	105	2,255	2,021
A	66	44	1,820	1,643
M	30	28	1,477	1,447
J	12	17	1,079	1,230
J	10	14	1,000	1,146
A	18	25	1,255	1,398
S	57	41	1,756	1,613
O	134	61	2,127	1,785
N	233	84	2,367	1,924
D	350	135	2,544	2,130

Soma 22,547 20,730

$\bar{X} = 1,878\ 9$, $\bar{Y} = 1,727\ 5$

$a = 0,649$, $b = 0,508\ 1$

$Y = 0,649\ x + 0,508\ 1$

$\sigma_x = 3,222\ x^{0,649}$

Na Fig. 9, acham-se representadas as médias em abscissas e os desvios-padrão em ordenadas. Note-se a estabilidade relativa das precipitações de outubro, novembro e dezembro (pontos observados abaixo da curva) e a instabilidade relativa das precipitações de janeiro, fevereiro e março (pontos acima da curva).

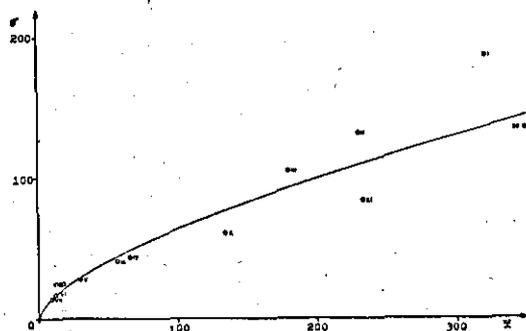


FIG. 9. Desvio-padrão e média das precipitações mensais em Morro Velho.

Tem-se, então,

$1 - a = 0,351$

e a transformação

$z = x^{0,351}$,

homogeneizando as variâncias, normalizará, possivelmente, a distribuição.

Os valores de $100z = 100 x^{0,351}$ nos diferentes meses foram os seguintes:

J	F	M	A	M	J	J	A	S	O	N	D	Média
757	675	619	435	330	239	224	276	413	558	677	781	498,7

Com estes dados, calculou-se:¹⁸

$$p_1 = 270,73 \quad p_2 = -7,67 \quad p_3 = -3,33 \quad p_4 = -5,17$$

$$p_5 = -0,90 \quad p_6 = 4,67$$

$$q_1 = -52,77 \quad q_2 = -6,35 \quad q_3 = -3,33 \quad q_4 = -13,57$$

$$q_5 = -12,07 \quad 6a_1^2 = 456 \ 476.$$

Como antes, quase toda a variação entre os meses (99,3%) corre à conta da onda anual, como se pode ver pelo teste de completicidade:

	G.L.	S Q.	Q.M.	F
Onda anual	2	456 476	228 238	650,4***
Desvio da onda anual	9	3 158	351	
	11	459 634		

Os valores de x , calculados por meio da função inversa de $z = x^{0,351}$, encontram-se, a seguir, em confronto com os valores observados.

QUADRO 6. Precipitações mensais em Morro Velho, Minas Gerais

Mês	Calculadas (onda anual)	Observadas
J	332	320
F	262	230
M	156	180
A	71	66
M	27	30
J	12	12
J	10	10
A	21	18
S	55	57
O	130	134
N	235	233
D	223	350

Na Fig. 10, como no Quadro 6, pode-se ver que o ajustamento foi bastante satisfatório.

CORRELAÇÃO ENTRE OS DADOS E UMA SUA HARMÔNICA

Pode-se mostrar (Apêndice, ítem 12) que o coeficiente de correlação entre os dados e sua harmônica de frequência m/N é

$$R_m = \sqrt{\frac{Na_m^2/2}{S(x_i - \bar{x})^2}}$$

¹⁸ A transformação $z = x^{0,351}$ deveria ser aplicada às precipitações mensais em cada ano; seria feita a análise harmônica das médias mensais dos valores transformados. Na falta dos dados originais, a transformação é feita aqui nos valores médios das precipitações mensais, à guisa de exemplo de aplicação.

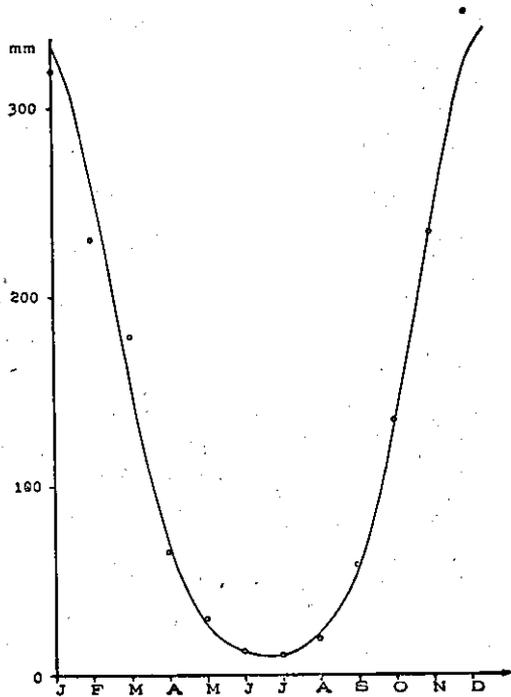


FIG. 10. Precipitações mensais em Morro Velho. Onda anual.

donde

$$\frac{N}{2} a_m^2 = R_m^2 \sum_{i=0}^{N-1} (x_i - \bar{x})^2.$$

Conclui-se que $\frac{N}{2} a_m^2$, que é a parte da variação entre os dados devida à harmônica de frequência m/N , é igual à parte daquela variação devida à correlação entre os dados e a harmônica de ordem m .

Se N é par e $m = N/2$, tem-se

$$R_{N/2} = \sqrt{\frac{Np_{N/2}^2}{S(x_i - \bar{x})^2}} = \sqrt{\frac{Na_{N/2}^2}{S(x_i - \bar{x})^2}}$$

$$Na_{N/2}^2 = R_{N/2}^2 S(x_i - \bar{x})^2,$$

de modo que a conclusão vale, também, para este caso excepcional.

Notar-se-á que R_m varia entre zero e 1, não podendo ser negativo.

SIGNIFICADO DO QUADRADO DA AMPLITUDE

Notando-se que

$$\sum_{i=0}^{N-1} (x_i - \bar{x}) \cos mri = \sum_{i=0}^{N-1} x_i \cos mri,$$

sendo $\bar{x} = \frac{1}{N} \sum_{i=0}^{N-1} x_i$ a média aritmética dos x_i , e

$$\sum_{i=0}^{N-1} (x_i - \bar{x}) \operatorname{sen} mri = \sum_{i=0}^{N-1} x_i \operatorname{sen} mri,$$

e substituindo em

$$a_m^2 = p_m^2 + q_m^2$$

p_m e q_m por seus valores

$$p_m = \frac{2}{N} \sum_{i=0}^{N-1} x_i \cos mri$$

$$q_m = \frac{2}{N} \sum_{i=0}^{N-1} x_i \operatorname{sen} mri$$

onde $r = 2\pi/N$ radianos ou $360/N$ graus e x_i representa, agora, os desvios em relação à média aritmética dos dados originais, pode-se exprimir o quadrado da amplitude da harmônica de ordem m na forma (Apêndice, item 13)

$$a_m^2 = \frac{4}{N^2} \sum_{i=0}^{N-1} x_i^2 \sum_{k=0}^{N-1} r_k \cos mrk,$$

onde

$$r_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k}}{\sum_{i=0}^{N-1} x_i^2},$$

convencionando-se que $x_{i+k} = x_{i+k-N}$ se $i+k > N-1$. Verifica-se, em particular, $r_0 = 1$.

Considerando a sucessão $x_0, x_1, x_2, \dots, x_{n-1}$ como sucessão cíclica, ou circular, x_0 seguindo-se a x_{n-1} , o autor sugere denominar-se r_k *coeficiente de correlação cíclica de ordem k*, ou *coeficiente de correlação circular de ordem k*, que difere do coeficiente de correlação serial¹⁹ por exprimir-se sempre em função do mesmo número N de produtos, qualquer que seja k .

Notando-se que

$$P_m = \frac{2}{N} \sum_{k=0}^{N-1} r_k \cos mrk$$

é a componente p da harmônica de frequência m/N dos coeficientes de correlação cíclica

$$r_0, r_1, r_2, \dots, r_{N-1},$$

¹⁹ Ver Kendall (1955) e Conrad e Pollak (1950), § 14.1. Kendall designa por coeficiente de autocorrelação o parâmetro do qual o coeficiente de correlação serial é uma estimativa. Outros autores, como Brooks e Carruthers (1953), designam o coeficiente de correlação serial por coeficiente de correlação retardada (*lag correlation*).

tem-se

$$a_m^2 = \frac{\sum_{i=0}^{N-1} x_i^2}{N/2} P_m,$$

ou seja

$$\frac{N}{2} a_m^2 = P_m \sum_{i=0}^{N-1} x_i^2.$$

Resulta que a componente p da harmônica de frequência m/N dos coeficientes de correlação cíclica é igual à fração da variação entre os dados que corre à conta da harmônica de ordem m :

$$P_m = \frac{N}{2} a_m^2 / \sum_{i=0}^{N-1} x_i^2.$$

Verifica-se, portanto, a dupla desigualdade

$$0 \leq P_m \leq 1,$$

o extremo inferior sendo assumido quando $a_m = 0$, isto é, na ausência da harmônica de frequência m/N , e o extremo superior sendo assumido quando toda a variação entre os dados corre à conta da harmônica de ordem m .

Comparando a expressão acima com a fórmula

$$\frac{N}{2} a_m^2 = R_{x,m}^2 \sum_{i=0}^{N-1} x_i^2,$$

parágrafo anterior, conclui-se que P_m é igual ao quadrado do coeficiente de correlação entre os dados originais e sua harmônica de frequência m/N .

Pode-se verificar que as componentes q das harmônicas dos coeficientes de correlação cíclica são todas nulas:

$$Q_m = \frac{2}{N} \sum_{k=0}^{N-1} r_k \operatorname{sen} mrk = 0 \quad (\text{Apêndice, item 14})$$

Conclui-se que a harmônica de frequência m/N dos coeficientes de correlação cíclica,

$$r_0, r_1, r_2, \dots, r_{N-1},$$

é

$$P_m \cos mri,$$

a comparar com a harmônica correspondente dos dados originais,

$$p_m \cos mri + q_m \operatorname{sen} mri = a_m \operatorname{sen} (mri + A_m) = a_m \cos [mri - (90 - A_m)],$$

que tem o mesmo período e apresenta uma defasagem de $90 - A_m$ graus.

No caso particular em que N é par e $m = N/2$, tem-se

$$N a_{N/2}^2 = P_{N/2} \sum_{i=0}^{N-1} x_i^2;$$

ainda neste caso, portanto, $P_{N/2}$ é igual à fração da variação entre os dados que corre à conta da harmônica de frequência $m/N = 1/2$. Conclui-se que

$$P_1 + P_2 + \dots + P_{I(N/2)} = 1.$$

Estes resultados, que, no referente à análise harmônica, parecem ser uma contribuição original, não só constituem uma introdução natural ao método de Fuhrich de análise periodográfica, a que se fará referência em próximo parágrafo, como podem servir de base à determinação das componentes harmônicas na ordem decrescente de suas amplitudes. Aplicando-se o método, determinar-se-ão, apenas, as componentes harmônicas cujas amplitudes sejam relevantes, como é usual em relação às ondas não harmônicas no método de Fuhrich. Compreender-se-á bem a importância desta observação quando se considerar o caso em que o número de harmônicas é considerável, como ocorre na análise harmônica de um elemento climático nas 73 pentadas no ano (36 componentes harmônicas).

Para duas componentes harmônicas quaisquer, de frequências m/N e n/N , tem-se

$$P_m/P_n = a_m^2/a_n^2,$$

de modo que, se a_m for maior que a_n , P_m será muito maior que P_n .

Continuando o processo, obtendo-se os coeficientes de correlação cíclica $r_k^{(s)}$ dos coeficientes de correlação cíclica $r_0, r_1, r_2, \dots, r_{N-1}$ e suas componentes $p, P_1^{(s)}, P_2^{(s)}, \dots, P_{I(N/2)}^{(s)}$, obtém-se, após as interações,

$$P_m^{(s)}/P_n^{(s)} = a_m^{2s}/a_n^{2s}$$

Como

$$P_1^{(s)} + P_2^{(s)} + \dots + P_{I(N/2)}^{(s)} = 1,$$

tem-se, portanto, se m é a componente harmônica de maior amplitude,

$$\lim_{s \rightarrow \infty} P_m^{(s)} = 1,$$

$$\lim_{s \rightarrow \infty} P_n^{(s)} = 0, \quad n \neq m.$$

A harmônica de maior amplitude predomina, assim, cada vez mais, obtendo-se, no limite,

$$r_i^{(\infty)} = \cos mri.$$

De

$$\sum_{i=0}^{N-1} x_i = 0 \quad \text{vem} \quad \sum_{i=0}^{N-1} x_i = 0,$$

donde se tira

$$\sum_{i=0}^{N-1} r_i = 0$$

e, em geral,

$$\sum_{i=0}^{N-1} r_i^{(s)} = 0.$$

Tem-se

$$\lim_{s \rightarrow \infty} \sum_{i=0}^{N-1} [r_i^{(s)}]^s = \sum_{i=0}^{N-1} \cos^s mri = N/2,$$

de modo que a variância de $r^{(s)}$ tende a $1/2$:

$$\lim_{s \rightarrow \infty} \sigma_s^2 = \lim_{s \rightarrow \infty} \sum_{i=0}^{N-1} [r_i^{(s)}]^s / N = 1/2.$$

A diferença

$$\Delta = 1/2 - \sigma_s^2$$

permite, assim, verificar, quando convenientemente pequena, que se alcançou um predomínio suficiente da onda de maior amplitude sobre as demais.²⁰

Se duas ou mais harmônicas tiverem igual amplitude, maior que as das demais harmônicas, do que se poderá suspeitar pela lentidão da convergência dos coeficientes de correlação cíclica a uma onda senoidal simples, recorrer-se-á ao mesmo processo aplicado no método de Fuhrich, substituindo-se aos $r^{(s)}$ suas diferenças sucessivas.

$$r_0^{(s)} - r_1^{(s)}, r_1^{(s)} - r_2^{(s)}, \dots, r_{N-2}^{(s)} - r_{N-1}^{(s)}, r_{N-1}^{(s)} - r_0^{(s)}.$$

Tem-se, por exemplo, para $s = 1$,

$$r_i = P_1 \cos ri + P_2 \cos 2ri + \dots + P_m \cos mri + \dots + P_n \cos nri + \dots,$$

$$r_i - r_{i+1} = 2 \left[\sin \frac{r}{2} P_1 \sin r(i+1/2) + \sin \frac{2r}{2} P_2 \sin 2r(i+1/2) + \dots + \sin \frac{nr}{2} P_n \sin nr(i+1/2) + \dots \right],$$

cujas componentes harmônicas de frequência m/N e n/N são

$$2 \sin \frac{mr}{2} P_m \sin mr(i+1/2)$$

e

$$2 \sin \frac{nr}{2} P_n \sin nr(i+1/2).$$

Se $a_m = a_n$ e, portanto, $P_m = P_n$, as amplitudes dessas harmônicas não são iguais, mas proporcionais

a $\sin \frac{mr}{2}$ e $\sin \frac{nr}{2}$. Como $mr/2$ e $nr/2$ pertencem ao primeiro quadrante, $\sin \frac{mr}{2}$ será maior que $\sin \frac{nr}{2}$ se $m > n$, predominando, pois, das ondas de igual amplitude, aquela de maior frequência.

²⁰ Conrad e Pollak (1950). Se N é par e $m = N/2$ é a frequência da onda de maior amplitude,

$$\lim_{s \rightarrow \infty} \sigma_s^2 = 1 \text{ e } \Delta = 1 - \sigma_s^2.$$

ANÁLISE HARMÔNICA E CORRELAÇÃO CÍCLICA

No caso particular da análise harmônica, onde os períodos possíveis são bem determinados, é suficiente examinar o andamento dos coeficientes de correlação cíclica, como explicado a seguir, tornando-se, assim, dispensável a repetição do processo.

Tem-se $r_0 = 1$ e, em virtude da simetria dos coeficientes de correlação cíclica, isto é, de $r_{N-i} = r_i$, resulta que, se a sucessão

$$r_0, r_1, \dots, r_{I(N/2)},$$

onde $I(N/2)$ é o maior inteiro que não supera $N/2$, for monótona decrescente, $r_{I(N/2)}$ será, certamente, um mínimo. É certa, pois, a existência de um mínimo absoluto, situado, quando muito, em $i = I(N/2)$ e o mínimo é, certamente, negativo, pois $Sr_i = 0$.

Marcando-se os coeficientes de correlação cíclica, r_i , em ordenadas e i em abscissas, ter-se-á traçado o correlograma.

Em virtude da preponderância da onda de maior amplitude no andamento do correlograma, o valor de i correspondente ao mínimo permitirá identificar a frequência da onda de maior amplitude: assim, se r_i for o mínimo, deve-se ter

$$\cos mri = -1, \quad mri = \pi,$$

donde

$$m = N/2i,$$

indicando $m/N = 1/2i$ como a frequência da onda presumivelmente de maior amplitude. A correção da indicação será confirmada pela concordância dos sinais dos coeficientes de correlação cíclica com os valores correspondentes de $\cos mri$. A necessidade desta concordância de sinais resulta do significado da amplitude relativamente à correlação entre os dados e sua harmônica de frequência correspondente. Resulta que $P_m > 0$ é proporcional ao coeficiente de correlação entre os coeficientes de correlação cíclica e sua harmônica de frequência m/N . Esta correlação, necessariamente positiva, deve ser a maior possível se a harmônica for a de maior amplitude, donde a necessidade da concordância de sinais.

Determinada a frequência da onda de maior amplitude, volta-se aos dados originais para determinar a amplitude e o ângulo fase.²¹

Não é necessário recorrer aos dados originais para, eliminada a onda predominante, determinar as de-

mais. A eliminação da onda predominante pode ser feita nos coeficientes de correlação cíclica, calculando-se novos coeficientes r'_k , pela fórmula

$$r'_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk}{\sum_{i=0}^{N-1} x_i^2 - \frac{N}{2} a_m^2},$$

x_i representando o desvio em relação à média aritmética. Note-se que o denominador e o numerador de r'_k se obtêm a partir dos termos correspondentes de r_k , subtraindo a parte da variação devida à onda predominante e, respectivamente, a parte da covariação cíclica de ordem k devida à mesma onda (Apêndice, item 15).

Determinada a segunda onda em importância, pela análise correlográfica dos r'_k , sua eliminação pode ser feita nos r'_k , determinando-se novos coeficientes de correlação cíclica

$$r''_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk - \frac{N}{2} a_n^2 \cos nrk}{\sum_{i=0}^{N-1} x_i^2 - \frac{N}{2} a_m^2 - \frac{N}{2} a_n^2}$$

e, assim, sucessivamente, até que se tenha eliminado da variação entre os dados toda a parte considerada relevante por um teste de completicidade.

Exemplificar-se-á a aplicação do método na análise harmônica das precipitações mensais em Pelotas (médias de 52 anos), que figuram na última coluna do Quadro 1:

J	F	M	A	M	J	
103,9	132,5	106,0	101,4	101,9	117,5	
J	A	S	O	N	D	Soma
106,1	117,8	128,1	99,5	78,6	82,1	1 275,4

Estes dados são transcritos em 12 setores correspondentes de dois círculos concêntricos de cartolina, o menor dos quais pode girar em torno do centro (Fig. 11).²²

²² Se N é grande; não é fácil desenhar os N setores em cada círculo de modo que os raios nos dois círculos se correspondam em cada posição da seta, como adiante indicado. É preferível então, transcrever os dados em coluna em duas tiras de papel quadriculado, em uma das quais, após o último dado, se inscrevem novamente os $I(N/2) + 1$ primeiros. Desenhando uma seta no alto da primeira tira e designando por 0, 1, 2, as linhas sucessivas na segunda, faz-se corresponder a seta com a linha zero e se multiplicam os números correspondentes; anota-se a soma dos produtos, faz-se corresponder a seta com a linha 1 e se repete o processo e, assim, sucessivamente.

²¹ Ver no Apêndice, itens 16, 17 e 18 como se pode determinar as amplitudes e os ângulos fase sem voltar aos dados originais.

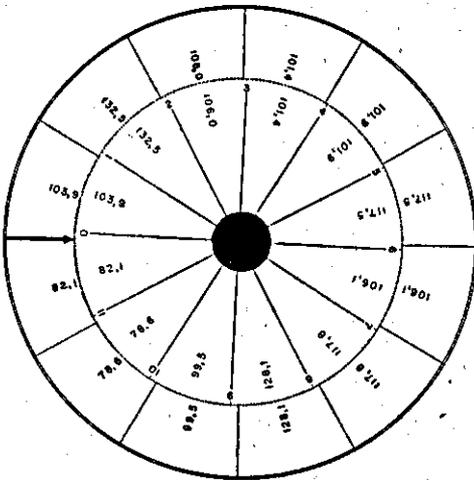


FIG. 11. Dispositivo para o cálculo dos coeficientes de correlação cíclica.

Fazendo corresponder o ponto zero do círculo menor com a seta assinalada no círculo maior, a soma dos produtos dos números correspondentes é

$$\sum_{i=0}^{11} x_i^2 = 138\ 421,56.$$

Fazendo corresponder o ponto 1 do círculo menor com a seta assinalada no círculo maior, anota-se a soma dos produtos

$$\sum_{i=0}^{11} x_i x_{i+1} = 136\ 471,47.$$

Passa-se, em seguida, ao ponto 2 do círculo menor e, assim, sucessivamente, obtendo-se

$$\sum_{i=0}^{11} x_i^2 = 138\ 421,56 \quad \sum_{i=0}^{11} x_i x_{i+1} = 136\ 471,47$$

$$\sum_{i=0}^{11} x_i x_{i+2} = 134\ 412,80 \quad \sum_{i=0}^{11} x_i x_{i+3} = 134\ 094,51$$

$$\sum_{i=0}^{11} x_i x_{i+4} = 135\ 253,04 \quad \sum_{i=0}^{11} x_i x_{i+5} = 135\ 923,70$$

$$\sum_{i=0}^{11} x_i x_{i+6} = 135\ 912,56 \quad \sum_{i=0}^{11} x_i x_{i+7} = 135\ 923,70$$

$$\sum_{i=0}^{11} x_i x_{i+8} = 135\ 253,04 \quad \sum_{i=0}^{11} x_i x_{i+9} = 134\ 094,51$$

$$\sum_{i=0}^{11} x_i x_{i+10} = 134\ 412,80 \quad \sum_{i=0}^{11} x_i x_{i+11} = 136\ 471,47$$

Notar-se-á que os resultados se repetem simetricamente a partir $\sum_{i=0}^{11} x_i x_{i+1(N/2)}$.

Esta propriedade é geral, como demonstrado no Apêndice.

Para verificar os resultados, deve-se observar que o total das somas dos produtos (incluindo a soma dos quadrados) deve conferir com o quadrado da soma, $1\ 275,4^2 = 1\ 626\ 645,16$. Tem-se, efetivamente:

$$138\ 421,56 + 2(136\ 471,47 + 134\ 412,80 + 134\ 094,51 + 135\ 253,04 + 135\ 923,70) + 135\ 912,56 = 1\ 626\ 645,16.$$

A soma dos quadrados dos desvios da média e as somas dos produtos dos desvios da média são obtidas, subtraindo a correção $(\sum_{i=0}^{N-1} x_i)^2 / N$, no caso $1\ 275,4^2 / 12 = 135\ 553,76$, dos resultados acima.

Tem-se, portanto;

$$\begin{aligned} S(x_i - \bar{x})^2 &= 2\ 867,80 \quad S(x_i - \bar{x})(x_{i+1} - \bar{x}) = 917,71 \\ S(x_i - \bar{x})(x_{i+2} - \bar{x}) &= -1\ 140,96 \quad S(x_i - \bar{x})(x_{i+3} - \bar{x}) = -1\ 459,25 \\ S(x_i - \bar{x})(x_{i+4} - \bar{x}) &= -300,72 \quad S(x_i - \bar{x})(x_{i+5} - \bar{x}) = 369,94 \\ S(x_i - \bar{x})(x_{i+6} - \bar{x}) &= 358,80 \quad S(x_i - \bar{x})(x_{i+7} - \bar{x}) = 369,94 \\ S(x_i - \bar{x})(x_{i+8} - \bar{x}) &= -300,72 \quad S(x_i - \bar{x})(x_{i+9} - \bar{x}) = -1\ 459,25 \\ S(x_i - \bar{x})(x_{i+10} - \bar{x}) &= -1\ 140,96 \quad S(x_i - \bar{x})(x_{i+11} - \bar{x}) = 917,71, \end{aligned}$$

cujas somas, salvo erros de aproximação, deve ser igual a zero.

Dividindo estes números por $S(x_i - \bar{x})^2 = 2\ 867,80$, obtêm-se os coeficientes de correlação cíclica

$$\begin{aligned} r_0 &= 1 \quad r_1 = \frac{917,71}{2\ 867,80} = 0,320\ 0 \\ r_2 &= \frac{-1\ 140,96}{2\ 867,80} = -0,397\ 9 \\ r_3 &= \frac{-1\ 459,25}{2\ 867,80} = -0,508\ 8 \\ r_4 &= \frac{-300,72}{2\ 867,80} = -0,104\ 9 \quad r_5 = \frac{369,94}{2\ 867,80} = 0,129\ 0 \\ r_6 &= \frac{358,80}{2\ 867,80} = 0,125\ 1 \quad r_7 = r_5 = 0,129\ 0 \\ r_8 &= r_4 = -0,104\ 9 \quad r_9 = r_3 = -0,508\ 8 \\ r_{10} &= r_2 = -0,397\ 9 \quad r_{11} = r_1 = 0,320\ 0. \end{aligned}$$

A soma desses coeficientes, salvo erros de aproximação, deve ser igual a zero.

A análise do correlograma (Fig. 12), ou o simples exame dos coeficientes de correlação cíclica, indica

r_3 como mínimo absoluto, logo a frequência da onda de maior amplitude deve ser $m/N = 1/2i = 1/6$, que é a frequência da onda semestral.

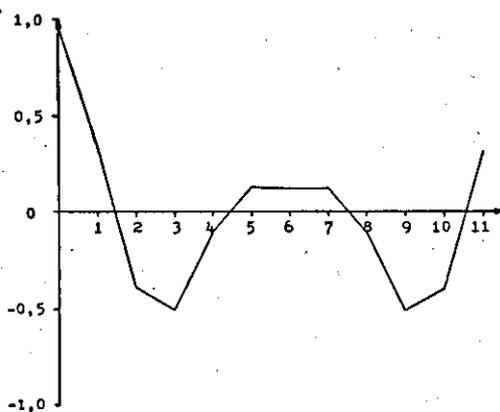


FIG. 12. Correlograma das precipitações mensais em Pelotas.

Comparando os coeficientes de correlação cíclica com os valores correspondentes de $\cos mri = \cos 60i$ (exprimindo o arco em graus), verifica-se a perfeita concordância dos sinais respectivos, o que confirma ser, de fato, a onda semestral a harmônica de maior amplitude:

i	0	1	2	3	4	5
r_i	1,000 0	0,320 0	-0,397 9	-0,508 8	-0,104 9	0,129 0
$\cos 60i$	1,000	0,500	-0,500	-1,000	-0,500	0,500
i	6	7	8	9	10	11
r_i	0,125 1	0,129 0	-0,104 9	-0,508 8	-0,397 9	0,320 0
$\cos 60i$	1,000	0,500	-0,500	-1,000	-0,500	0,500

Note-se que, apesar de ser $\Delta = 1/2 - \sigma_1^2 = 0,342 2$ relativamente considerável, foi possível identificar a harmônica de maior amplitude pelo simples exame dos coeficientes de correlação cíclica, sem necessidade de iterar o processo.

Embora se possa calcular a amplitude a partir dos coeficientes de correlação cíclica, deve-se voltar aos dados originais para determinar o ângulo fase.²³

Tem-se $a_2 = 15,69$ (dividindo por 52 a amplitude da onda semestral dos totais mensais nos 52 anos) e $A_2 = 16^\circ 27'$.

A componente p da onda semestral dos coeficientes de correlação cíclica pode obter-se, multiplicando por $2/N = 1/6$ a soma dos produtos dos valores correspondentes de r_i e de $\cos mri$, resultando

$$P_2 = \frac{1}{6} \times 3,094 5;$$

²³ Ver apêndice, itens 16, 17 e 18.

onde

$$a_2^2 = \frac{Sx_i^2}{6} P_2 = \frac{2 867,80 \times 3,094 5}{36} = 246,511 2$$

e $a_2 = 15,70$, a comparar com o valor $a_2 = 15,69$, obtido diretamente.

Para eliminar a onda semestral, aplica-se a fórmula

$$r'_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk}{\sum_{i=0}^{N-1} x_i^2 - \frac{N}{2} a_m^2}$$

obtendo-se

$$r'_0 = 1 \quad r'_1 = \frac{917,71 - 733,22}{2 867,80 - 1 476,45} = \frac{179,49}{1 391,35} = 0,129 0$$

$$r'_2 = \frac{-402,74}{1 391,35} = -0,289 5 \quad r'_3 = \frac{17,20}{1 391,35} = 0,012 4$$

$$r'_4 = \frac{437,50}{1 391,35} = 0,314 4 \quad r'_5 = \frac{-368,28}{1 391,35} = -0,264 7$$

$$r'_6 = \frac{-1 117,65}{1 391,35} = -0,803 3 \quad r'_7 = r'_8 = -0,264 7$$

$$r'_8 = r'_9 = 0,314 4 \quad r'_{10} = r'_{11} = 0,012 4$$

$$r'_{10} = r'_{11} = -0,289 5 \quad r'_{11} = r'_{10} = 0,129 0.$$

Salvo erros de aproximação, a soma dos r'_i deve ser nula. Note-se a conveniência de conservar o numerador e o denominador dos r'_i , tal como se fez com os r_i , pois serão usados no prosseguimento da análise.

O mínimo absoluto (Fig. 13) é $r'_2 = -0,289 5$, indicando $m/N = 1/2i = 1/4$ como a frequência da componente harmônica que se segue em importância à onda semestral.

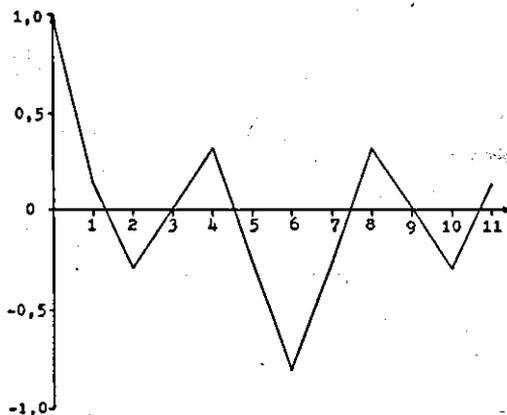


FIG. 13. Correlograma, após a eliminação da onda semestral.

O primeiro ponto de anulamento é $i = 1,3$, o que corresponde a $m = N/4i = 2,3$, ou seja, a $m/N = 1/5,2$. A falta de concordância desta indicação parece decorrer, não só da imprecisão do correlograma mas, também, como se verá adiante, da pequena diferença entre as amplitudes das ondas quadrimestral e anual.

Tem-se $\Delta = 1/2 - 0,1820 = 0,3180$.

A concordância com os sinais de $\cos mri = \cos 90i$ é perfeita, como se pode ver adiante.

i	0	1	2	3	4	5
r_i'	1,000 0	0,129 0	-0,289 5	0,012 4	0,314 4	-0,264 7
$\cos 90i$	1	0	-1	0	1	0
i	6	7	8	9	10	11
r_i'	-0,803 3	-0,264 7	0,314 4	0,012 4	-0,289 5	0,129 0
$\cos 90i$	-1	0	1	0	-1	0

Tem-se

$$P_3 = \frac{1}{6} \times 3,011 1, \quad a_3^2 = \frac{1 \ 391,35 \times 3,011 1}{36} = 116,374 8,$$

donde $a_3 = 10,79$, a comparar com $a_3 = 10,75$, determinada com os dados originais.

Eliminando a onda quadrimestral, tem-se

$$r_0'' = 1 \quad r_1'' = \frac{179,49 - 0}{1 \ 391,35 - 693,92} = \frac{179,49}{697,43} = 0,257 4$$

$$r_2'' = \frac{291,18}{697,43} = 0,417 5 \quad r_3'' = \frac{17,20}{697,43} = 0,024 7$$

$$r_4'' = \frac{-256,42}{697,43} = -0,367 7 \quad r_5'' = \frac{-368,23}{697,43} = -0,528 1$$

$$r_6'' = \frac{-423,73}{697,43} = -0,607 6 \quad r_7'' = r_8'' = -0,528 1$$

$$r_8'' = r_9'' = -0,367 7 \quad r_{10}'' = r_{11}'' = 0,257 4$$

O mínimo absoluto (Fig. 14) é $r_6'' = -0,607 6$, logo, $m/N = 1/2i = 1/12$ deve ser a frequência da terceira onda em importância (onda anual), o que é confirmado pela concordância dos sinais de r_i'' e de $\cos 30i$:

i	0	1	2	3	4	5
r_i''	1,000 0	0,257 4	0,417 5	0,024 7	-0,367 7	-0,528 1
$\cos 30i$	1,000	0,866	0,500	0	-0,500	-0,866
i	6	7	8	9	10	11
r_i''	-0,607 6	-0,528 1	-0,367 7	0,024 7	0,417 5	0,257 4
$\cos 30i$	-1,000	-0,866	-0,500	0	0,500	0,866

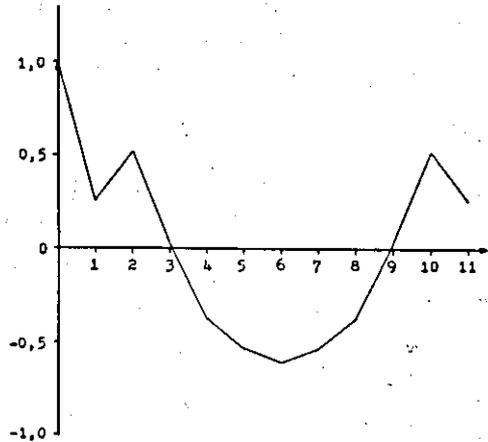


FIG. 14. Correlograma, após a eliminação da onda quadrimestral.

Quanto ao anulamento, tem-se $i = 3,1$, donde $m = N/4i = 0,97 \cong 1$, ou seja $m/N = 1/12,4 \cong 1/12$.

Tem-se

$$\Delta = 1/2 - 0,223 3 = 0,276 7, \quad P_1 = \frac{1}{6} \times 3,753 3,$$

$$a_1^2 = \frac{697,43 \times 3,753 3}{36} = 72,712 7$$

donde $a_1 = 8,53$, a comparar com $a_1 = 8,52$, calculada com os dados originais.

Eliminando a onda anual, tem-se

$$r_0''' = 1 \quad r_1''' = \frac{179,49 - 377,39}{697,43 - 435,78} =$$

$$\frac{-197,90}{261,65} = -0,756 4$$

$$r_2''' = \frac{73,29}{261,65} = 0,280 1 \quad r_3''' = \frac{17,20}{261,65} = 0,065 7$$

$$r_4''' = \frac{-33,53}{261,65} = -0,128 3 \quad r_5''' = \frac{9,11}{261,65} = 0,034 8$$

$$r_6''' = \frac{12,05}{261,65} = 0,046 1 \quad r_7''' = r_8''' = 0,034 8$$

$$r_9''' = r_{10}''' = -0,128 3 \quad r_{11}''' = r_{12}''' = 0,065 7$$

$$r_{13}''' = r_{14}''' = 0,280 1 \quad r_{15}''' = r_{16}''' = -0,756 4$$

O mínimo absoluto sendo $r_1''' = -0,756 4$ (Fig. 15), presumir-se-ia que a onda mais importante, após a remoção das três primeiras, seria a de frequência

$m/N = 1/2i = 1/2$ (onda bimestral)²⁴; o confronto com os sinais dos cossenos mostra, entretanto, que não é esta, mas sim a de frequência $m/N = 1/2,4$ a que se segue em importância:

i	0	1	2	3	4	5
r_i''''	1,000 0	-0,756 4	0,280 1	0,065 7	-0,147 3	0,034 8
$\cos 180i$	1	-1	1	-1	1	-1
$\cos 150i$	1,000	-0,866	0,500	0	-0,500	0,866

i	6	7	8	9	10	11
r_i''''	0,016 1	0,034 8	-0,147 3	0,065 7	0,280 1	-0,756 4
$\cos 180i$	1	-1	1	-1	1	-1
$\cos 150i$	-1,000	0,866	-0,500	0	0,500	-0,866

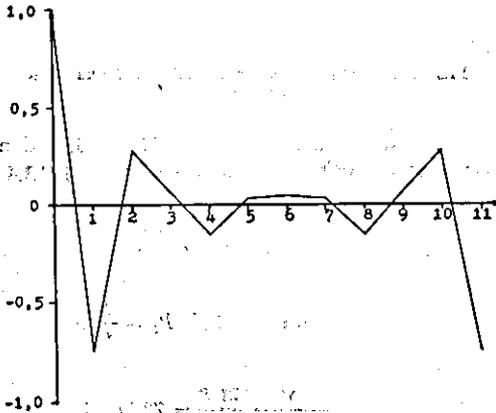


FIG. 15. Correlograma, após a eliminação da onda anual.

Em seis casos, como se vê, o sinal de r_i'''' não concorda com o de $\cos 180i$, ao passo que em apenas um caso ($r_6'''' = 0,016 1$; $\cos 900^\circ = -1$) não concorda com o sinal de $\cos 150i$. Esta discordância, mesmo em um único caso, constitui, evidentemente, um indicio da pequena relevância da quinta onda e, com maior razão, das demais. É neste ponto que cabe fazer o teste de completicidade, que revela a desnecessidade de prosseguir o processo, se se deseja determinar apenas as harmônicas que contribuem substancialmente na variação entre os dados.

Notar-se-á que a indicação dada pelo mínimo absoluto, sobre a frequência da onda mais importante, não falhou, ainda esta vez: é que só se dispõe de r_i'''' correspondentes a valores inteiros de i , ao passo que para a quinta harmônica o mínimo deveria corresponder a $i = 1,2$.

²⁴ Quando, como no caso presente, N é par e $m = N/2$, a importância da harmônica de frequência m é dada pelo dobro do quadrado da amplitude, visto como, neste caso

$$\sum_{i=0}^{N-1} (x_i - \bar{x})^2 = \frac{N}{2} (a_1^2 + a_2^2 + a_3^2 + \dots + a_{(N-2)/2}^2 + N a_{N/2}^2)$$

No caso presente, verifica-se que $2a_5^2 < a_6^2$, de modo que é esta última a que se segue em importância.

Quanto ao anulamento, tem-se $i = 0,6$, $m/N = 1/4i = 1/2,4$, indicando corretamente a onda de maior amplitude.

ITERAÇÃO DO PROCESSO

Substituindo dúvida sobre a harmônica mais relevante após a remoção das três primeiras, cabe iterar o processo, determinando-se os coeficientes de correlação cíclica de segunda ordem, isto é, os coeficientes de correlação cíclica $r_i''''^{(2)}$ dos coeficientes de correlação cíclica r_i'''' , usando-se o mesmo dispositivo antes descrito.

Obtém-se

i	0	1	2	3	4	5
$r_i''''^{(2)}$	1,000 0	-0,816 9	0,399 8	-0,023 0	-0,134 1	0,110 1
$\cos 150i$	1,000	-0,866	0,500	0	-0,500	0,866

i	6	7	8	9	10	11
$r_i''''^{(2)}$	-0,071 9	0,110 1	-0,134 1	-0,023 0	0,399 8	-0,816 9
$\cos 150i$	-1,000	0,866	-0,500	0	0,500	-0,866

A concordância com os sinais de $\cos 150i$ é agora perfeita, indicando a quinta harmônica como a onda mais importante após a remoção das três primeiras.

Embora seja desnecessário prosseguir o processo, foram determinados os coeficientes de correlação cíclica de terceira, quarta e quinta ordens, para que se possa ter uma idéia da rapidez da convergência a uma onda senoidal simples:

i	0	1	2	3	4	5
r_i''''	1,000 0	-0,756 4	0,280 1	0,035 7	-0,147 3	0,034 8
$r_i''''^{(2)}$	1,000 0	-0,816 9	0,399 8	-0,023 0	-0,134 1	0,110 1
$r_i''''^{(3)}$	1,000 0	-0,861 7	0,523 2	-0,153 7	-0,103 5	0,232 5
$r_i''''^{(4)}$	1,000 0	-0,891 7	0,603 5	-0,228 1	-0,126 3	0,372 4
$r_i''''^{(5)}$	1,000 0	-0,891 1	0,593 8	-0,188 2	-0,216 7	0,512 7
$\cos 150i$	1,000	-0,866	0,500	0	-0,500	0,866

i	6	7	8	9	10	11
r_i''''	0,016 1	0,034 8	-0,147 3	0,065 7	0,280 1	-0,756 4
$r_i''''^{(2)}$	-0,071 9	0,110 1	-0,134 1	-0,023 0	0,399 8	-0,816 9
$r_i''''^{(3)}$	-0,263 5	0,232 5	-0,103 5	-0,153 7	0,523 2	-0,861 7
$r_i''''^{(4)}$	-0,459 7	0,372 4	-0,126 3	-0,228 1	0,603 5	-0,891 7
$r_i''''^{(5)}$	-0,620 9	0,512 7	-0,216 7	-0,188 2	0,593 8	-0,891 1
$\cos 150i$	-1,000	0,866	-0,500	0	0,500	-0,866

Tem-se $\Delta_1 = 1/2 - \sigma^2 = 0,303 5$, $\Delta_2 = 0,273 3$, $\Delta_3 = 0,226 6$, $\Delta_4 = 0,171 4$ e $\Delta_5 = 0,135 9$.

DUAS HARMÔNICAS COM IGUAL AMPLITUDE

Se duas harmônicas de frequências m/N e n/N , têm igual amplitude, $a_m = a_n = a$, as harmônicas

correspondentes dos coeficientes de correlação cíclica são

$$P_m \cos mri = P \cos mri \text{ e } P_n \cos nri = P \cos nri, \quad m > n.$$

Ora,

$$\cos mri + \cos nri = 2 \cos \frac{m+n}{2} ri \cos \frac{m-n}{2} ri$$

anula-se para $\frac{m+n}{2} ri = 90^\circ$ e $\frac{m-n}{2} ri = 90^\circ$.

O primeiro anulamento corresponde, pois, a

$$\frac{m+n}{2} ri = 90^\circ \text{ ou } i = 180 / (m+n) r = N / 2(m+n).$$

Pode-se verificar que $\cos mri + \cos nri$ é decrescente neste ponto, passando pois, de valores positivos a valores negativos. Efetivamente, fazendo

$$i = \frac{\pi}{(m+n)r} + \Delta i, \quad \text{em radianos,}$$

tem-se, sendo Δi convenientemente pequeno em valor absoluto,

$$\begin{aligned} \cos mri + \cos nri &\approx \cos \frac{\pi m}{m+n} + \cos \frac{\pi n}{m+n} - \\ &- r \Delta i \left(m \operatorname{sen} \frac{\pi m}{m+n} + n \operatorname{sen} \frac{\pi n}{m+n} \right), \end{aligned}$$

visto como, para pequenos valores de Δi ,

$$\begin{aligned} \operatorname{sen} mr \Delta i &\approx mr \Delta i, \quad \operatorname{sen} nr \Delta i \approx nr \Delta i, \\ \cos nr \Delta i &\approx \cos nr \Delta i \approx 1. \end{aligned}$$

Mas, sendo

$$\frac{\pi m}{m+n} + \frac{\pi n}{m+n} = \pi,$$

vem a ser

$$\operatorname{sen} \frac{\pi m}{m+n} = \operatorname{sen} \frac{\pi n}{m+n}, \quad \cos \frac{\pi m}{m+n} = -\cos \frac{\pi n}{m+n} < 0,$$

pois $\frac{\pi n}{m+n} < \pi/2$, por ser $n < m$. Resulta que

$$\cos mri + \cos nri \approx -r \Delta i (m+n) \operatorname{sen} \frac{\pi n}{m+n}$$

tem, na vizinhança do primeiro ponto de anulamento, sinal contrário ao de Δi , sendo pois, positiva à esquerda e negativa à direita; em outras palavras, $\cos mri + \cos nri$ é decrescente no primeiro ponto de anulamento, passando de valores positivos a valores negativos.

Presumir-se-ia que o mínimo correspondesse ao dobro do valor relativo ao primeiro ponto de anulamento, como ocorre normalmente, isto é, a

$\frac{2\pi}{(m+n)r} = N / (m+n)$. Pode-se verificar, entretanto, aplicando o mesmo processo acima, que nesse último ponto $\cos mri + \cos nri$ é crescente. O mínimo deve estar situado, pois, entre $i = 180 / (m+n)r = N / 2(m+n)$, correspondente ao primeiro

anulamento, e $i = 360 / (m+n)r = N / (m+n)$ que, em casos normais, indicaria o mínimo.

Conclui-se que, no caso de duas harmônicas predominantes, com iguais amplitudes, o primeiro ponto de anulamento deve indicar $1/4i = (m+n)/2N$ como a frequência da onda predominante, e o mínimo absoluto indicaria $1/2i > (m+n)/2N$, mais próxima da maior frequência.

Como ilustração, depois de eliminadas as ondas semestral e quadrimestral, obtendo-se

$$\begin{aligned} r_0'' &= 1 & r_1'' &= \frac{179,49}{697,43} & r_2'' &= \frac{291,18}{697,43} & r_3'' &= \frac{17,20}{697,43} \\ r_4'' &= \frac{-256,42}{697,43} & r_5'' &= \frac{-368,28}{697,43} & r_6'' &= \frac{-423,73}{697,43} \text{ etc.} \end{aligned}$$

introduzir-se-á uma onda quadrimestral com a mesma amplitude da onda anual. Tem-se $a_3'' = a_1'' = 72,630$ e os novos coeficientes de correlação cíclica são obtidos dos coeficientes correspondentes acima transcritos, somando $N a_3'' / 2 = 435,78$ ao denominador e $435,78 \cos mrk = 435,78 \cos 90k$ ao denominador.

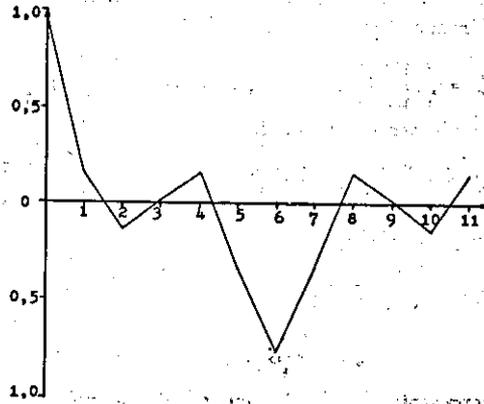


FIG. 16. Duas ondas com igual amplitude.

Como se vê no correlograma (Fig. 16), o primeiro ponto de anulamento, $i = 1,5$, indica $1/4i = 1/6$ como a frequência da onda mais importante, ao passo que o mínimo, $i = 2$, aponta para $m/N = 1/2i = 1/4$, que coincide com a onda predominante de maior frequência.

O confronto com os sinais dos cossenos correspondentes à onda quadrimestral é perfeito, como se pode ver adiante:

i	0	1	2	3	4	5
(r_1'')	1,000 0	0,158 4	-0,127 6	0,015 2	0,158 3	-0,325 0
$\cos 90i$	1	0	-1	0	1	0
i	6	7	8	9	10	11
(r_1'')	-0,758 5	-0,325 0	0,158 3	0,015 2	-0,127 6	0,158 4
$\cos 90i$	-1	0	1	0	-1	0

Ainda neste caso, como se vê, o mínimo absoluto deu indicação correta, localizando, entre as duas ondas de maior amplitude, a de maior frequência.

PROCESSOS PRÁTICOS DE CÁLCULO

O processo de cálculo ilustrado na parte inicial do presente trabalho (Marcha da Análise) pode ser simplificado, tirando-se partido de que os valores dos cossenos e dos senos, necessários para o cálculo dos p e dos q, se repetem, com o mesmo sinal ou com sinal contrário. O processo sistemático de cálculo que daí resulta, podendo servir para a preparação de um programa para computador, encontra-se em Brooks e Carruthers (1953), § 18.3, para N = 12 e N = 24.

Um segundo processo, baseado no novo conceito de coeficiente de correlação cíclica, é também ilustrado neste trabalho (Análise Harmônica e Correlação Cíclica), permitindo a determinação das diversas harmônicas na ordem de sua importância e a eliminação de cada uma delas sem voltar aos dados originais (Apêndice, itens 16, 17 e 18).

Um terceiro processo pode ser elaborado, a partir da fórmula (Apêndice, item 13):

$$a_m^s = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + 2 \sum_{i < j} x_i x_j \cos mr(j-i) \right],$$

que pode ser escrita na forma

$$a_m^s = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + 2 \sum_{k=1}^{N-1} \left(\sum_{i=0}^{N-k-1} x_i x_{i+k} \right) \cos mrk \right],$$

ou, ainda, fazendo

$$2 \sum_{i=0}^{N-k-1} x_i x_{i+k} = R_k, \quad \sum_{i=0}^{N-1} x_i^2 = R_0$$

$$a_m^s = \frac{4}{N^2} \sum_{k=0}^{N-1} R_k \cos mrk,$$

x_i representando desvio em relação à média aritmética.

Tem-se

$$\sum_{k=0}^{N-1} R_k = 0,$$

o que permite verificar a correção dos cálculos.

Pode-se proceder com os R_K da mesma forma indicada para os R_K correspondentes aos coeficientes de correlação cíclica, mas parece não ser possível eliminar as harmônicas sucessivas sem recurso aos dados originais.

Os dois últimos processos podem ser, também, programados para computador, sendo indicados no caso em que o número N de dados é considerável, desejando-se determinar, apenas, as harmônicas mais importantes.

OUTROS MÉTODOS DE ANÁLISE PERIODOGRÁFICA

A análise harmônica é apenas um dos métodos de análise dos fenômenos periódicos, aliás o mais simples e mais bem sistematizado.

Muitas vezes, o período fundamental é claramente indicado pelas condições físicas do fenômeno, como ocorre com o curso anual de observações meteorológicas. Pode acontecer, entretanto, que ondas senoidais cujos períodos não são submúltiplos do ano interfiram no curso de um fenômeno periódico.

Suponhamos, por exemplo, que uma onda anual se combine com uma onda de período igual a nove meses; uma análise harmônica, com período fundamental de doze meses, não poderia revelar a existência da segunda onda, de período igual a nove meses. Neste caso, a determinação da segunda onda e de suas possíveis harmônicas, poderia ser feita submetendo à análise harmônica os dados originais reagrupados em períodos de nove meses.

Teríamos assim,

J ₁	F ₁	M ₁	A ₁	M ₁	J ₁	J ₁	A ₁	S ₁
O ₁	N ₁	D ₁	J ₂	F ₂	M ₂	A ₂	M ₂	J ₂
J ₂	A ₂	S ₂	O ₂	N ₂	D ₂	J ₃	F ₃	M ₃
A ₃	M ₃	J ₃	J ₃	A ₃	S ₃	O ₃	N ₃	D ₃
.
.

T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	T ₉
----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------

onde as letras indicam os meses e os índices anos consecutivos. O número total de meses deve ser múltiplo comum de 9 e 12, isto é, de 36.

Os totais das nove colunas são, então, submetidos à análise harmônica, com os valores apropriados dos cossenos e dos senos, para o cálculo das amplitudes e dos ângulos fase.

O problema complica-se se ocorrem ondas de período desconhecido. Métodos especiais foram elaborados para a resolução do problema (ver Conrad & Pollak 1950, Brooks & Carruthers 1953), entre os quais o do periodograma de Schuster. Reagrupando cada vez os dados originais em períodos arbitrários, determina-se para cada período e seus submúltiplos as amplitudes das ondas respectivas. Representando num gráfico os períodos em abscissas e as amplitudes (ou seus quadrados) em ordenadas, obtém-se o periodograma. Os máximos relativos

da curva assim obtida indicam que na sua vizinhança devem existir ondas reais, que uma análise minuciosa na região correspondente permite localizar.

A melhor solução obtida até agora para o problema é o método desenvolvido por Fuhrich (Conrad & Pollak 1950), baseado nos coeficientes de autocorrelação de ordem 1, 2, N - k. Repetindo o processo, a amplitude da onda mais conspícua tende a 1 e as demais tendem a zero, conservando-se o período de cada onda. Determinado o período da onda mais conspícua, volta-se a considerar a sucessão original para determinar a amplitude e a fase. Elimina-se, em seguida, o efeito da onda assim determinada da sucessão original e se repete o processo, determinando-se a onda seguinte, na ordem decrescente das amplitudes e, assim, sucessivamente.

Se as duas ondas mais conspícuas têm a mesma amplitude, a convergência do processo é retardada. Neste caso, substitui-se à última sucessão obtida $x'_1, x'_2, x'_3, \dots, x'_n$, a nova sucessão $x'_2 - x'_1, x'_3 - x'_2, \dots, x'_n - x'_{n-1}$ os períodos $\frac{360^\circ}{f_1}, \frac{360^\circ}{f_2}, \dots$ das diversas ondas não se alteram, mas as amplitudes das ondas mais conspícuas, de igual amplitude, tendem a $\text{sen } \frac{f_2}{2}, \text{sen } \frac{f_3}{2}$ etc., e as demais tendem a zero.²⁵

Embora o método de Fuhrich tenha relegado a segundo plano os métodos baseados na análise harmônica, nos casos mais gerais em que se desconhece o período fundamental do fenômeno, pode-se aplicar a análise harmônica para facilitar a construção das sucessões derivadas da sucessão original (Conrad & Pollak 1950, § 14.6).

Baseando-se nas idéias de Von Mises, Fuhrich nega possibilidade de submeter as amplitudes determinadas pelo seu método a testes de significância, limitando-se a submetê-las a testes de completicidade, cujos detalhes o leitor interessado encontrará em Conrad e Pollak (1950), § 14.2.1.

NOVAS PERSPECTIVAS

Em tudo que precede, o fenômeno periódico foi considerado como resultado da interferência de ondas senoidais permanentes (com amplitude, período e fase constante) e de perturbações de caráter aleatório.

A situação complica-se se as características das ondas senoidais variam, se algumas delas se amortecem

²⁵ f é a frequência angular da onda, $f = 360/t$, onde t é o período.

ou desaparecem (ondas transitórias), outras têm período crescente, outras, ainda, sofrem variações de fase.

Em 1927, Yule chamou a atenção para os fenômenos auto-regressivos, em que uma variação casual, uma vez ocorrida, se incorpora ao sistema, comportando-se daí em diante, como se fôsse um elemento sistemático.

Os grandes êxitos alcançados no passado na superação das dificuldades que se apresentavam no estudo dos fenômenos periódicos, não deixam dúvida de que as novas dificuldades serão, também, superadas.

AGRADECIMENTOS

O autor agradece ao Eng.^o Agrônomo João Baptista da Silva, que tomou a si grande parte do trabalho de cálculo, ao Eng.^o Agrônomo João Gilberto Corrêa da Silva, que fez o teste de homogeneidade das variâncias, aos Srs. Nestor Carvalho de Oliveira e Luiz Carlos Aranha Prietch, que auxiliaram nos cálculos e fizeram a cópia datilográfica, aos Srs. Eduardo Fortuna e Gilberto Galhardi, desenhistas, ao Sr. Rubens Fernandes Castanheira, fotógrafo, ao estudante Luiz Augusto Regis do Amaral, Agronomando de 1967, por sua colaboração na execução dos cálculos, e ao Dr. Kent B. Tyler, do IRI, pela revisão do sumário em inglês.

REFERÊNCIAS

- Amaral, E. & Mota, F.S. 1956. Normais e variabilidade relativa das precipitações mensais. Bol. Téc. 13, Inst. Agron. do Sul, Pelotas.
- Blackwood, O.H., Herron, W.B. & Kelly, W.C. 1962. Física na escola secundária. Min. Educ. Cultura, Rio de Janeiro.
- Bliss, C.I. 1958. Periodic regression in biology and climatology. Bol. 615, Connecticut Agr. Exp. Sta., New Haven.
- Brooks, C.E.P. & Carruthers, N. 1953. Handbook of statistical methods in meteorology. Meteorol. Office, London.
- Cochran, W.G. & Cox, G.M. 1960. Experimental designs 2nd ed. John Wiley and Sons, New York and London.
- Conrad, V. & Pollak, L.W. 1950. Methods in climatology. 2nd ed. Harvard Univ. Press, Cambridge, Massachusetts.
- Deming, W.E. 1964. Statistical adjustment of data. Dovere Publ., New York.
- Duncan, D. B. 1955. Multiple range and multiple F tests. Biometrics 11:1-42.
- Fisher, R.A. 1941. Statistical methods for research workers. 8th ed. Oliver and Boyd, London.
- Fisher, R.A. 1950a. On the similarity of the distributions found for the test of significance in harmonic analysis, and in Stevens, problem in geometrical probability. In Contributions to mathematical statistics. John Wiley and Sons, New York and London.
- Fisher, R. A. 1950b. Tests of significance in harmonic analysis. In Contributions to mathematical statistics John Wiley and Sons, New York and London.

- Fisher, R.A. & Yates, F. 1949. Statistical tables for biological, agricultural and medical research. 3th ed. Oliver and Boyd, London.
- Gomes, F.P. 1963. Curso de estatística experimental, 2.^a ed. Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, São Paulo.
- Kendall, M.G. 1955. The advanced theory of statistics. Vol. 2. 3rd ed., 2nd print. Charles Griffin, London.
- Memória, J.M.P. 1955a. Análise estatística das precipitações pluviométricas mensais de Morro Velho. Bol. de Agricultura n.º 5 da Secretaria da Agricultura do Estado de Minas Gerais. (Separata)
- Memória, J.M.P. 1955b. Análise estatística das precipitações pluviométricas mensais de Morro Velho. Bol. de Agricultura n.º 6 da Secretaria da Agricultura do Estado de Minas Gerais. (Separata)
- Snedecor, G.W. 1957. Statistical methods. 5th ed. Iowa State College Press, Ames, Iowa.
- Van Valkenburgh; Nooger and Neville, Inc. 1960. Eletricidade básica. Vol. 3. Livraria Freitas Bastos, Rio de Janeiro.
- Waugh, A. E. -. Elementos de estatística. Editôra Globo, Pôrto Alegre. (Sem data)

HARMONIC ANALYSIS

Abstract

The theoretical foundation of the harmonic analysis and its calculation procedure are presented.

An example is studied, in which the variance was not independent of the mean, and thus the data were not normally distributed. In spite of this, the contrasts necessary to the harmonic analysis have shown homogeneity of variance. This has permitted the direct processing of the data (monthly precipitations at Pelotas, Rio Grande do Sul, Brazil, 1900/51).

In another example, the data were submitted to an earlier transformation, to normalize, presumably, the distribution, then the transformed data were analysed (monthly precipitations at Morro Velho, Minas Gerais, Brazil, 1855/1951).

The stability of the harmonic components of the monthly precipitations at Pelotas was tested by the variance analysis method. The waves with periods of six, four and twelve months were significant in regard to their variation between years. The second and the third harmonics were predominant in the yearly course of the monthly precipitations (maxima at February and September). The completeness test reveals that 91.2% of the variation of the monthly precipitation means was accounted for those three waves.

Since yearly data concerning the monthly precipitations at Morro Velho were not available to the author, it was impossible to submit them to a stability test. Practically all the yearly variation of the monthly precipitation means (99,3%) resulted from the yearly wave, which was to be anticipated in a continental tropical climate.

Two original contributions are made in this paper: the establishment of the distribution of the difference between the squares of two amplitudes and the introduction of the concept of cyclic correlation coefficient. The latter is not to be confounded with the known serial correlation coefficient. That new concept is applied in the adaptation of the Fuhrich's periodographical method to the harmonic analysis, permitting the isolation of the harmonics in the order of their importance, as measured by their amplitudes.

APÊNDICE

Fundamentos trigonométricos

1. A soma dos cossenos de uma sucessão de arcos em progressão aritmética é dada pela fórmula

$$\cos a + \cos(a+r) + \cos(a+2r) + \dots + \cos[a + (n-1)r] = \frac{\operatorname{sen} \frac{nr}{2} \cos\left(a + \frac{n-1}{2}r\right)}{\operatorname{sen} \frac{r}{2}}$$

Tem-se, efetivamente,

$$\begin{aligned} \operatorname{sen}\left(a+ir + \frac{r}{2}\right) - \operatorname{sen}\left(a+ir - \frac{r}{2}\right) &= \\ &= 2 \operatorname{sen} \frac{r}{2} \cos(a+ir) \end{aligned}$$

Fazendo, sucessivamente, $i = 0, 1, 2, \dots, n-1$, tem-se

$$\operatorname{sen}\left(a + \frac{r}{2}\right) - \operatorname{sen}\left(a - \frac{r}{2}\right) = 2 \operatorname{sen} \frac{r}{2} \cos a$$

$$\operatorname{sen}\left(a + \frac{3r}{2}\right) - \operatorname{sen}\left(a + \frac{r}{2}\right) = 2 \operatorname{sen} \frac{r}{2} \cos(a+r)$$

$$\operatorname{sen}\left(a + \frac{5r}{2}\right) - \operatorname{sen}\left(a + \frac{3r}{2}\right) = 2 \operatorname{sen} \frac{r}{2} \cos(a+2r)$$

$$\begin{aligned} \operatorname{sen}\left(a + \frac{2n-1}{2}r\right) - \operatorname{sen}\left(a + \frac{2n-3}{2}r\right) &= \\ &= 2 \operatorname{sen} \frac{r}{2} \cos\left[a + (n-1)r\right]. \end{aligned}$$

Somando membro a membro, vem

$$\begin{aligned} n\left(a + \frac{2n-1}{2}r\right) - \operatorname{sen}\left(a - \frac{r}{2}\right) &= 2 \operatorname{sen} \frac{r}{2} \left\{ \cos a + \right. \\ &+ \cos(a+r) + \dots + \cos[a + (n-1)r] \left. \right\}. \end{aligned}$$

Mas

$$\begin{aligned} \operatorname{sen}\left(a + \frac{2n-1}{2}r\right) - \operatorname{sen}\left(a - \frac{r}{2}\right) &= \\ &= 2 \operatorname{sen} \frac{nr}{2} \cos\left(a + \frac{n-1}{2}r\right), \end{aligned}$$

donde, substituindo,

$$\begin{aligned} \cos a + \cos(a+r) + \cos(a+2r) + \dots + \cos[a + \\ + (n-1)r] &= \frac{\operatorname{sen} \frac{nr}{2} \cos\left(a + \frac{n-1}{2}r\right)}{\operatorname{sen} \frac{r}{2}}. \end{aligned}$$

2. Análogamente, a soma dos senos de uma sucessão de arcos em progressão aritmética é dada pela fórmula

$$\begin{aligned} \operatorname{sen} a + \operatorname{sen}(a+r) + \operatorname{sen}(a+2r) + \dots + \operatorname{sen}[a + \\ + (n-1)r] &= \frac{\operatorname{sen} \frac{nr}{2} \operatorname{sen}\left(a + \frac{n-1}{2}r\right)}{\operatorname{sen} \frac{r}{2}}, \end{aligned}$$

como se pode verificar, substituindo na fórmula precedente a por $\frac{\pi}{2} - a$ e r por $-r$.

3. A soma dos quadrados dos cossenos de uma sucessão de arcos em progressão aritmética é dada pela fórmula

$$\begin{aligned} \cos^2 a + \cos^2(a+r) + \cos^2(a+2r) + \dots + \cos^2[a + \\ + (n-1)r] &= \frac{n}{2} + \frac{\operatorname{sen} nr \cos[2a + (n-1)r]}{2 \operatorname{sen} r} \end{aligned}$$

Realmente, de

$$\cos^2 a - \operatorname{sen}^2 a = \cos 2a,$$

vem

$$2 \cos^2 a - 1 = \cos 2a,$$

$$2 \cos^2(a+r) - 1 = \cos(2a+2r)$$

$$2 \cos^2(a+2r) - 1 = \cos(2a+4r)$$

$$2 \cos^2[a + (n-1)r] - 1 = \cos[2a + 2(n-1)r].$$

Somando membro a membro e aplicando a fórmula da soma dos cossenos de uma sucessão de arcos em progressão aritmética, obtém-se a fórmula desejada.

Se $r = k\pi$, k inteiro, a soma dos quadrados dos cossenos reduz-se a $n \cos^2 a$.

4. Semelhantemente, a soma dos quadrados dos senos de uma sucessão de arcos em progressão aritmética é dada pela fórmula

$$\begin{aligned} \operatorname{sen}^2 a + \operatorname{sen}^2(a+r) + \operatorname{sen}^2(a+2r) + \dots + \operatorname{sen}^2[a + \\ + (n-1)r] &= \frac{n}{2} - \frac{\operatorname{sen} nr \cos[2a + (n-1)r]}{2 \operatorname{sen} r}, \end{aligned}$$

visto como a soma dos quadrados dos senos de n arcos mais a soma dos quadrados dos cossenos desses arcos é igual a n .

Se $r = k\pi$, k inteiro, a soma dos quadrados dos senos reduz-se a $m \operatorname{sen}^2 a$.

5. Com métodos semelhantes, obtém-se as fórmulas:

$$\begin{aligned} & \text{sen } a \cos a + \text{sen } (a + r) \cos (a + r) + \dots + \\ & + \text{sen } [a + (n - 1)r] \cos [a + (n - 1)r] = \\ & = \frac{\text{sen } nr \text{sen } [2a + (n - 1)r]}{2 \text{sen } r}, \end{aligned}$$

que se reduz a $n \text{sen } a \cos a$ se $r = k\pi$;

$$\begin{aligned} & \cos a \cos a + \cos (a + mr) \cos (a + pr) + \dots + \\ & + \cos [a + (n - 1)mr] \cos [a + (n - 1)pr] = \\ & = \frac{\text{sen } \frac{n(m+p)r}{2} \cos \left[2a + \frac{(n-1)(m+p)r}{2} \right]}{2 \text{sen } \frac{(m+p)r}{2}} + \\ & + \frac{\text{sen } \frac{n(m-p)r}{2} \cos \frac{(n-1)(m-p)r}{2}}{2 \text{sen } \frac{(m-p)r}{2}} \end{aligned}$$

sendo $|m| \neq |p|$.

Se $(m + p)r = 2k\pi$, a primeira fração reduz-se a $\frac{n}{2} \cos 2a$ e, se $(m - p)r = 2k\pi$, a segunda fração reduz-se a $\frac{n}{2}$.

Se $p = -m$, a soma dos produtos dos cossenos é igual a

$$\begin{aligned} & \frac{n}{2} \cos 2a + \frac{\text{sen } nmr \cos (n - 1)mr}{2 \text{sen } mr} \\ & \text{sen } a \text{sen } a + \text{sen } (a + mr) \text{sen } (a + pr) + \dots + \\ & + \text{sen } [a + (n - 1)mr] \text{sen } [a + (n - 1)pr] = \\ & = \frac{\text{sen } \frac{n(m-p)r}{2} \cos \frac{(n-1)(m-p)r}{2}}{2 \text{sen } \frac{(m-p)r}{2}} - \\ & - \frac{\text{sen } \frac{n(m+p)r}{2} \cos \left[2a + \frac{(n-1)(m+p)r}{2} \right]}{2 \text{sen } \frac{(m+p)r}{2}} \end{aligned}$$

As fórmulas para os casos particulares, $(m + p)r = 2k\pi$ e $(m - p)r = 2k\pi$, podem ser obtidas das fórmulas correspondentes para o produto dos cossenos, substituindo a por $\frac{\pi}{2} - a$ e r por $-r$.

$$\begin{aligned} & \text{sen } a \cos a + \text{sen } (a + mr) \cos (a + pr) + \dots + \\ & + \text{sen } [a + (n - 1)mr] \cos [a + (n - 1)pr] = \\ & = \frac{\text{sen } \frac{n(m+p)r}{2} \text{sen } \left[2a + \frac{(n-1)(m+p)r}{2} \right]}{2 \text{sen } \frac{(m+p)r}{2}} + \\ & + \frac{\text{sen } \frac{n(m-p)r}{2} \text{sen } \frac{(n-1)(m-p)r}{2}}{2 \text{sen } \frac{(m-p)r}{2}}, \end{aligned}$$

sendo $|m| \neq |p|$.

Se $(m + p)r = 2k\pi$, a primeira fração reduz-se a $\frac{n}{2} \text{sen } 2a$ e, se $(m - p)r = 2k\pi$, a segunda fração reduz-se a zero.

6. No caso que interessa especialmente, no presente trabalho;

$$r = \frac{2\pi}{n}, \quad 0 < m \leq \frac{n}{2}, \quad 0 < p < \frac{n}{2}, \quad m \neq p$$

inteiros; as fórmulas reduzem-se, então, às seguintes:

$$\begin{aligned} & \cos a + \cos (a + mr) + \dots + \cos [a + (n - 1)mr] = 0; \\ & \text{sen } a + \text{sen } (a + mr) + \dots + \text{sen } [a + (n - 1)mr] = 0; \\ & \cos^2 a + \cos^2 (a + pr) + \dots + \cos^2 [a + (n - 1)pr] = \\ & = \frac{n}{2}; \end{aligned}$$

$$\begin{aligned} & \text{sen}^2 a + \text{sen}^2 (a + pr) + \dots + \text{sen}^2 [a + (n - 1)pr] = \\ & = \frac{n}{2}; \end{aligned}$$

$$\begin{aligned} & \cos^2 a + \cos^2 (a + mr) + \dots + \cos^2 [a + (n - 1)mr] = \\ & = n \cos^2 a, \quad m = \frac{n}{2}; \end{aligned}$$

$$\begin{aligned} & \cos a \text{sen } a + \cos (a + pr) \text{sen } (a + pr) + \dots + \\ & + \cos [a + (n - 1)pr] \text{sen } [a + (n - 1)pr] = 0; \end{aligned}$$

$$\begin{aligned} & \cos a \cos a + \cos (a + mr) \cos (a + pr) + \dots + \\ & + \cos [a + (n - 1)mr] \cos [a + (n - 1)pr] = 0; \end{aligned}$$

$$\begin{aligned} & \text{sen } a \text{sen } a + \text{sen } (a + mr) \text{sen } (a + pr) + \dots + \\ & + \text{sen } [a + (n - 1)mr] \text{sen } [a + (n - 1)pr] = 0; \end{aligned}$$

$$\begin{aligned} & \cos a \text{sen } a + \cos (a + mr) \text{sen } (a + pr) + \dots + \\ & + \cos [a + (n - 1)mr] \text{sen } [a + (n - 1)pr] = 0. \end{aligned}$$

7. Resulta das fórmulas precedentes a ortogonalidade das $n - 1$ formas lineares

$$\begin{aligned} & x_1 \cos 0 + x_2 \cos r + x_3 \cos 2r + \dots + x_n \cos (n - 1)r \\ & x_1 \cos 0 + x_2 \cos 2r + x_3 \cos 4r + \dots + x_n \cos 2(n - 1)r \\ & x_1 \cos 0 + x_2 \cos 3r + x_3 \cos 6r + \dots + x_n \cos 3(n - 1)r \\ & \dots \dots \dots \end{aligned}$$

$$\begin{aligned} & x_1 \cos 0 + x_2 \cos I_n r + x_3 \cos 2I_n r + \dots + \\ & + x_n \cos I_n (n - 1)r \end{aligned}$$

$$\begin{aligned} & x_1 \text{sen } 0 + x_2 \text{sen } r + x_3 \text{sen } 2r + \dots + x_n \text{sen } (n - 1)r \\ & x_1 \text{sen } 0 + x_2 \text{sen } 2r + x_3 \text{sen } 4r + \dots + x_n \text{sen } 2(n - 1)r \end{aligned}$$

$$\begin{aligned} & x_1 \text{sen } 0 + x_2 \text{sen } 3r + x_3 \text{sen } 6r + \dots + x_n \text{sen } 3(n - 1)r \\ & \dots \dots \dots \end{aligned}$$

$$\begin{aligned} & x_1 \text{sen } 0 + x_2 \text{sen } I_{n-1} r + x_3 \text{sen } 2I_{n-1} r + \dots + \\ & + x_n \text{sen } I_{n-1} (n - 1)r, \end{aligned}$$

onde $r = \frac{2\pi}{n}$, $I_n = I(n/2)$ é o maior inteiro que não supera $n/2$ e I_{n-1} é o maior inteiro que não supera $(n - 1)/2$.

Se n é ímpar, $I_{n-1} = I_n = (n - 1)/2 < n/2$, e a norma de tôdas as $n - 1$ formas lineares (soma dos quadrados dos coeficientes) é igual a $n/2$.

Se n é par, $I_n = n/2$ e $I_{n-1} = (n/2) - 1$; neste caso, a norma da forma linear

$$x_1 \cos 0 + x_2 \cos I_n r + x_3 \cos 2I_n r + \dots + x_n \cos (n - 1)I_n r$$

é igual a n e a de tôdas as demais é igual a $n/2$.

Teste múltiplo de amplitudes

8. Segundo o critério de Newman-Keuls, quando diversas amplitudes são testadas, deve-se levar em conta que a probabilidade de determinado valor ser excedido por uma amplitude aumenta com o número de amplitudes menores do que a que está sendo testada. Deve-se pois, segundo este critério, manter o nível de 5% de probabilidade para cada teste, levando em conta a ordenação das amplitudes testadas.²⁶

O problema consiste em determinar o nível p a exigir para a significância de uma amplitude, de tal modo que o valor correspondente àquele nível seja excedido por uma, ou mais, de n amplitudes com probabilidade $\alpha = 0,05$.

Tem-se

$$1 - (1 - p)^n = 0,05, \\ (1 - p)^n = 0,95,$$

donde

$$\text{Log}(1 - p) = \frac{\text{Log } 0,95}{n}.$$

Sendo p pequeno; $\text{Log}(1 - p)$ não difere muito de $-p$; tem-se, portanto,

$$p \simeq \frac{-\text{Log } 0,95}{n}.$$

Mas, como visto anteriormente, o valor da amplitude correspondente ao nível de probabilidade p é

$$a = \sigma \sqrt{\frac{-\text{Log } p}{N}}.$$

Tem-se, portanto, aproximadamente

$$a \simeq 2 \sigma \sqrt{\frac{\text{Log}(n!) - \text{Log } 0,95}{N}},$$

ou

$$a \simeq 2 \sigma \sqrt{\frac{(\text{Log } 20 n)}{N}}$$

visto como $\text{Log } 0,95 = \text{Log}(1 - 0,05)$ pouco difere de $-0,05 = -1/20$.

Fazendo $\lambda_n = 2\sqrt{\text{Log}(20n)}$, o valor a comparar com a amplitude a testar é

$$a \simeq \lambda_n \sigma / \sqrt{N},$$

sendo λ_n dado no Quadro 3.

Como visto no texto do presente trabalho, o critério acima é discutível, por demasiado exigente, quando se testam amplitudes de ondas independentes,

²⁶ Como informam Brooks e Carruthers (1953), e também Fisher (1950b), foi Sir Gilbert Walker o primeiro a formular esta exigência.

como as que provêm de uma análise harmônica, ou mesmo quando se testa a raiz quadrada da diferença entre os quadrados de duas amplitudes independentes, como proposto no texto.

No caso presente, o critério de Duncan (1955) leva a tomar $\lambda_n = \lambda_1$ para o teste de cada amplitude ou da raiz quadrada do módulo da diferença entre os quadrados das amplitudes de duas ondas independentes, sem ter em conta a ordenação entre as amplitudes testadas.

Em casos mais gerais de análises periodográfica, em que as ondas não são independentes, o critério de Duncan (1955) exige que se tome em consideração a ordenação das amplitudes testadas, tal como ocorre no teste de Duncan (1955) de amplitude múltipla de n médias.

Nesse caso mais geral da análise periodográfica, entretanto, o problema dos testes de significância está por estudar, limitando-se Fuhrich a fazer testes de completicidade (Conrad & Pollak 1950).

Distribuição do módulo da diferença entre os quadrados de duas amplitudes

9. A distribuição conjunta dos quadrados de duas amplitudes é dada pela diferencial de segunda ordem

$$\frac{N^2}{16 \sigma^4} e^{-N(\sigma_1^2 + \sigma_2^2)/4 \sigma^2} d\sigma_1^2 d\sigma_2^2$$

Supondo $a_2 > a_1$ e fazendo

$$a_2^2 - a_1^2 = D,$$

fixando a_1^2 , de modo que D varie com a_1^2 , tem-se

$$a_2^2 = a_1^2 + D, \quad da_2^2 = dD$$

$$\frac{N^2}{16 \sigma^4} e^{-N(\sigma_1^2 + D)/4 \sigma^2} da_1^2 dD.$$

Integrando em relação a a_1^2 , de zero a $+\infty$, obtém-se a distribuição de D :

$$dP_D = \frac{N}{8 \sigma^2} e^{-ND/4 \sigma^2} dD,$$

cujas integral, estendida ao intervalo $0 < D < +\infty$ é igual a $1/2$.

A distribuição de $|D|$ é, portanto,

$$dP = \frac{N}{4 \sigma^2} e^{-N|D|/4 \sigma^2} d|D|, \quad 0 < |D| < +\infty.$$

Tal é a probabilidade condicionada de se obter $a_2^2 - a_1^2$ no intervalo $(D, D + dD)$, quando se sabe que $a_2 > a_1$, isto é que $D > 0$.

Se esta dedução é correta, conclui-se que a diferença entre dois χ^2 , cada um com dois graus de liberdade, se distribui também como χ^2 com dois graus de liberdade.

Distribuição de $R = a_2^2/a_1^2$

10. A distribuição do quadrado a_1^2 de uma amplitude é, como foi visto,

$$\frac{N}{4\sigma^2} e^{-Na_1^2/4\sigma^2} da_1^2.$$

Se a_2 é a amplitude de uma onda independente da primeira, procedente da análise harmônica de observações com a mesma variância σ^2 , a probabilidade conjunta de se obter a_1^2 e a_2^2 nos intervalos respectivos

$$(a_1^2, a_1^2 + da_1^2) \text{ e } (a_2^2, a_2^2 + da_2^2).$$

é, a menos de um infinitésimo de ordem superior à primeira em relação a $da_1^2 da_2^2$:

$$\frac{N^2}{16\sigma^4} e^{-N(a_1^2 + a_2^2)/4\sigma^2} da_1^2 da_2^2.$$

Fazendo $R = a_2^2/a_1^2$, donde $a_2^2 = Ra_1^2$, fixando a_1 e da_1 e deixando R variar com a_2^2 , tem-se

$$da_2^2 = a_1^2 dR.$$

A distribuição conjunta de a_1^2 e R é, então,

$$\frac{N^2 a_1^2}{16\sigma^4} e^{-Na_1^2(1+R)/4\sigma^2} da_1^2 dR.$$

Integrando em relação a a_1^2 , de zero a $+\infty$, obtém-se

$$dP = \frac{1}{(1+R)^2} dR, \quad 0 < R < +\infty.$$

R distribui-se, pois, como F com 2 e 2 graus de liberdade.

Teste de homogeneidade das variâncias residuais das três ondas: de 3 meses, de 12/5 meses e de 2 meses (51 g.l.). Teste de Bartlett²⁷

11. O teste de homogeneidade abaixo transcrito foi efetuado pelo Eng.º Agrônomo João Gilberto Corrêa da Silva.

Ondas	Q.N.(s²)	log s²
3 meses	1 19 579,640 4	4,201 80
	2 34 468,483 6	4,537 21
12/5 meses	3 29 827,440 2	4,474 62
	4 24 036,940 0	4,380 88
2 meses	5 28 242,656 1	4,450 91
Soma	8s² = 136 155,160 3	S log s² = 22,135 42

$$\bar{s}^2 = \frac{8s^2}{5} = \frac{136\ 155,160\ 3}{5} = 27\ 231,032\ 1;$$

$$\log \bar{s}^2 = 4,435\ 06$$

$$5 \log \bar{s}^2 = 22,175\ 30$$

$$S \log s^2 = 22,135\ 42$$

$$\text{Diferença} = 0,039\ 88$$

²⁷ Ver Snedecor (1957), Seção 10.20.

$X^2 = 2,3026 \times 51 \times \text{Diferença}, 4 \text{ G. L.}$

$$\text{Fator de correção} = C = 1 + \frac{5+1}{3 \times 5 \times 51} = 1,007\ 84$$

$$X^2 = 2,3026 \times 51 \times 0,039\ 08 = 4,683$$

X^2 (corrigido) = 4,683/1,007 84 = 4 647 com 4 G.L.

Este valor não é significativo ao nível de 5% de probabilidade.

Correlação entre os dados e uma sua harmônica

12. O coeficiente de correlação entre os dados e sua harmônica de frequência m/N é

$$R_m = \frac{Sx_i(p_m \cos mri + q_m \sin mri)}{\sqrt{S(x_i - \bar{x})^2 S(p_m \cos mri + q_m \sin mri)^2}}$$

Mas

$$Sx_i(p_m \cos mri + q_m \sin mri) = p_m Sx_i \cos mri + q_m Sx_i \sin mri = \frac{N}{2} p_m^2 + \frac{N}{2} q_m^2 = \frac{N}{2} a_m^2, \quad m < \frac{N}{2},$$

e

$$S(p_m \cos mri + q_m \sin mri)^2 = p_m^2 S \cos^2 mri + 2p_m q_m S \sin mri \cos mri + q_m^2 S \sin^2 mri = \frac{N}{2} a_m^2,$$

visto como

$$S \cos^2 mri = S \sin^2 mri = N/2, \quad m < \frac{N}{2}.$$

Resulta

$$R_m = \sqrt{\frac{Na_m^2/2}{S(x_i - \bar{x})^2}}, \quad \text{se } m < \frac{N}{2},$$

Se N é par e $m = N/2$, pode-se verificar que

$$R_m = \sqrt{\frac{Na_m^2}{S(x_i - \bar{x})^2}},$$

de modo que, ainda neste caso, o quadrado do coeficiente de correlação é igual à fração da variação total que corre à conta da harmônica de frequência m/N .

Significado do quadrado da amplitude

13. Notando-se que

$$\sum_{i=0}^{N-1} (x_i - \bar{x}) \cos mri = \sum_{i=0}^{N-1} x_i \cos mri,$$

onde \bar{x} é a média aritmética dos x_i , e, semelhantemente,

$$\sum_{i=0}^{N-1} (x_i - \bar{x}) \sin mri = \sum_{i=0}^{N-1} x_i \sin mri,$$

e substituindo em

$$a_m^2 = p_m^2 + q_m^2.$$

p_m e q_m por seus valores

$$p_m = \frac{2}{N} \sum_{i=0}^{N-1} x_i \cos mri$$

$$q_m = \frac{2}{N} \sum_{i=0}^{N-1} x_i \operatorname{sen} mri,$$

onde $r = 2\pi/N$ radianos ou $r = 360/N$ graus e x_i representa, agora, o desvio em relação à média aritmética dos dados originais, têm-se

$$a_m^2 = \frac{4}{N^2} \left(\sum_{i=0}^{N-1} x_i \cos mri + \sum_{i=0}^{N-1} x_i \operatorname{sen} mri \right) = \\ + \frac{4}{N^2} \left(\sum_{i=0}^{N-1} x_i^2 + 2 \sum x_i x_j \cos mri \cos mrj + \right. \\ \left. + 2 \sum x_i x_j \operatorname{sen} mri \operatorname{sen} mrj \right),$$

sendo $i < j$.

Tendo-se em conta que $\cos mri \cos mrj + \operatorname{sen} mri \operatorname{sen} mrj = \cos mr(i-j) = \cos mr(j-i)$,

vem

$$a_m^2 = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + \cos mr(j-i) x_i x_j + \right. \\ \left. + S \cos mr(i-j) x_i x_j \right],$$

que se pode escrever, por ser $r = 2\pi/N$ radianos,

$$a_m^2 = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + S \cos mr(j-i) x_i x_j + \right. \\ \left. + S \cos mr(i+N-j) x_i x_j \right].$$

Os termos em $x_i x_j$ que, no primeiro somatório, têm $j-i = k$ e os termos em $x_j x_i$ que, no segundo somatório, têm $i+N-j = k$, têm o mesmo coeficiente $\cos mrk$; reunindo-os, vem

$$a_m^2 = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + \sum_{k=1}^{N-1} \cos mrk \left(\sum_{i=0}^{N-k-1} x_i x_{i+k} + \right. \right. \\ \left. \left. + \sum_{i=N-k}^{N-1} x_i x_{i+k-N} \right) \right].$$

Convencionando-se representar x_{j+k-N} por x_{j+k} , sendo $j+k > N$, tem-se

$$a_m^2 = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 + \sum_{k=1}^{N-1} \cos mrk \left(\sum_{i=0}^{N-k-1} x_i x_{i+k} + \right. \right. \\ \left. \left. + \sum_{i=N-k}^{N-1} x_i x_{i+k} \right) \right],$$

donde

$$a_m^2 = \frac{4}{N^2} \left(\sum_{i=0}^{N-1} x_i^2 + \sum_{k=1}^{N-1} \cos mrk \sum_{i=0}^{N-1} x_i x_{i+k} \right),$$

que se pode escrever, por ser $\cos 0 = 1$,

$$a_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} \left(\cos mrk \sum_{i=0}^{N-1} x_i x_{i+k} \right),$$

ou, ainda,

$$a_m^2 = \frac{4}{N^2} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} r_k x_i^2 \cos mrk$$

onde

$$r_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k}}{\sum_{i=0}^{N-1} x_i^2}$$

é o coeficiente de correlação cíclica de ordem k dos dados originais, x_i representando, na fórmula acima o desvio em relação à média aritmética.

Voltando a representar por x_i os dados originais, tem-se, portanto,

$$r_k = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=0}^{N-1} (x_i - \bar{x})^2},$$

que se pode calcular mais facilmente pela fórmula

$$r_k = \frac{\sum_{i=0}^{N-1} x_i x_{i+k} - \sum_{i=0}^{N-1} x_i \bar{x}}{\sum_{i=0}^{N-1} x_i^2 - \sum_{i=0}^{N-1} x_i \bar{x}}, \quad r_0 = 1.$$

Notando-se que

$$P_m = \frac{2}{N} \sum_{i=0}^{N-1} r_k \cos mrk$$

é a componente p da harmônica de ordem m dos coeficientes de correlação cíclica, tem-se

$$a_m^2 = \frac{2}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})^2 P_m,$$

donde

$$P_m = \frac{2}{N} a_m^2 \sum_{i=0}^{N-1} (x_i - \bar{x})^2.$$

Conclui-se que a componente p da harmônica de ordem m dos coeficientes de correlação cíclica é igual à fração da variação dos dados que corre à conta da harmônica de ordem m .

Pode-se verificar que os coeficientes de correlação cíclica cuja soma das ordens é igual a N são iguais, isto é

$$r_k = r_{N-k}.$$

De fato, substituindo i por $i+N-k$, tem-se

$$S x_i x_{i+k} = S x_{i+N-k} x_{i+N} = S x_{i+N-k} x_i = \\ = S x_i x_{i+N-k},$$

donde

$$r_k = r_{N-k}.$$

No caso particular em que N é par e $m = N/2$, tem-se

$$p_{N/2} = \frac{1}{N} \sum_{i=0}^{N-1} x_i \cos 180i = \frac{1}{N} [(x_0 + x_2 + \dots + x_{N-2}) - (x_1 + x_3 + \dots + x_{N-1})],$$

donde

$$p_{N/2}^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} x_i^2 [(r_0 + r_2 + \dots + r_{N-2}) - (r_1 + r_3 + \dots + r_{N-1})],$$

ou

$$a_{N/2}^2 = \frac{1}{N} \sum_{i=0}^{N-1} x_i^2 P_{N/2},$$

sendo

$$P_{N/2} = \frac{1}{N} [(r_0 + r_2 + \dots + r_{N-2}) - (r_1 + r_3 + \dots + r_{N-1})]$$

a componente p da harmônica de ordem $N/2$ dos dados originais.

Ainda neste caso, portanto,

$$P_{N/2} = Na_{N/2}^2 \sum_{i=0}^{N-1} (x_i - \bar{x})^2$$

é a parte da variação entre os dados que corre à conta da harmônica de ordem $N/2$.

Componentes q das harmônicas dos coeficientes de correlação cíclica

14. Não é difícil verificar que as componentes q das harmônicas dos coeficientes de correlação cíclica são tôdas nulas. De fato, tem-se idênticamente:

$$\text{sen } mr(j-i) + \text{sen } mr(i-j) = 0$$

$$S x_i x_j \text{sen } mr(j-i) + S x_i x_j \text{sen } mr(i-j) = 0, i < j,$$

e, sendo

$$\text{sen } mr(i+N-j) = \text{sen } mr(i-j),$$

por ser $r = 2\pi/N$ radianos, vem

$$S x_i x_j \text{sen } mr(j-i) + S x_i x_j \text{sen } mr(i+N-j) = 0.$$

Os termos em $x_i x_j$ que, no primeiro somatório, têm $j-i = k$, isto é, $j = i+k$, e os termos em $x_i x_j$

que, no segundo somatório, têm $i+N-j = k$, ou $i = j+k-N$, têm o mesmo coeficiente, $\text{sen } mrk$; reunindo-os, têm-se

$$\left(\sum_{i=0}^{N-k-1} x_i x_{i+k} + \sum_{i=N-k}^{N-1} x_i x_{i+k-N} \right) \text{sen } mrk$$

que se pode escrever, convencionando que

$$x_{i+k} = x_{i+k-N} \text{ se } i+k > N-1,$$

$$\left(\sum_{i=0}^{N-k-1} x_i x_{i+k} + \sum_{i=N-k}^{N-1} x_i x_{i+k} \right) \text{sen } mrk$$

ou, ainda, mudando j em i no segundo somatório e somando êste ao primeiro,

$$\text{sen } mrk \sum_{i=0}^{N-1} x_i x_{i+k}.$$

Tem-se, portanto, k variado de zero a $N-1$,

$$\sum_{k=0}^{N-1} \left(\text{sen } mrk \sum_{i=0}^{N-1} x_i x_{i+k} \right) = 0,$$

donde

$$\sum_{k=0}^{N-1} r_k \text{sen } mrk = 0^{23}$$

e, sendo

$$Q_m = \frac{2}{N} \sum_{k=0}^{N-1} r_k \text{sen } mrk,$$

$$Q_m = 0,$$

qualquer que seja m .

$$\sum_{k=0}^{N-1} r_k \text{sen} \left(\frac{N}{2} - k \right) z = \text{sen} \frac{Nz}{2}.$$

Efetivamente, se N fôr ímpar,

$$\sum_{k=0}^{N-1} r_k \text{sen} \left(\frac{N}{2} - k \right) z = r_0 \text{sen} \frac{Nz}{2} + \sum_{k=1}^r r_k \text{sen} \left(\frac{N}{2} - k \right) z + r_{N-k} \text{sen} \left(\frac{N}{2} - N + k \right) z$$

onde $r = (N-1)/2$, e, sendo $r_{N-k} = r_k$ e $r_0 = 1$,

$$\sum_{k=0}^{N-1} r_k \text{sen} \left(\frac{N}{2} - k \right) z = \text{sen} \frac{Nz}{2} + \sum_{k=1}^r r_k \left| \text{sen} \left(\frac{N}{2} - k \right) z - \text{sen} \left(\frac{N}{2} - k \right) z \right| = \text{sen} \frac{Nz}{2};$$

Se N fôr par,

$$\sum_{k=0}^{N-1} r_k \text{sen} \left(\frac{N}{2} - k \right) z = r_0 \text{sen} \frac{Nz}{2} + r_{N/2} \text{sen} \left(\frac{N}{2} - \frac{N}{2} \right) z + \sum_{k=1}^r r_k \left| \text{sen} \left(\frac{N}{2} - k \right) z - \text{sen} \left(\frac{N}{2} - k \right) z \right| = \text{sen} \frac{Nz}{2}.$$

No caso particular em que Nz é múltiplo de 360° , vem

$$\sum_{k=0}^{N-1} r_k \text{sen } kz = 0.$$

²³ De modo mais geral, pode-se demonstrar a identidade

Eliminação da onda predominante

15. A eliminação da onda predominante pode ser feita nos coeficientes de correlação cíclica, sem necessidade de voltar aos dados originais.

Se m/N é a frequência da onda predominante, a eliminação dessa onda dos dados originais é feita pela fórmula

$$x'_i = x_i - p_m \cos mri - q_m \sin mri.$$

Elevando ao quadrado, vem

$$x_i'^2 = x_i^2 - 2p_m x_i \cos mri - 2q_m x_i \sin mri + p_m^2 \cos^2 mri + q_m^2 \sin^2 mri + 2p_m q_m \sin mri \cos mri.$$

Mas, supondo $m < N/2$

$$S x_i \cos mri = \frac{N}{2} p_m, \quad S x_i \sin mri = \frac{N}{2} q_m$$

$$S \cos^2 mri = N/2, \quad S \sin^2 mri = N/2,$$

$$2 S \sin mri \cos mri = S \sin 2mri = 0,$$

donde

$$S x_i'^2 = S x_i^2 - \frac{N}{2} a_m^2.$$

Analogamente,

$$x'_i x'_{i+k} = x_i x_{i+k} - p_m [x_i \cos mr(i+k) + x_{i+k} \cos mri] - q_m [x_i \sin mr(i+k) + x_{i+k} \sin mri] + p_m^2 \cos mr(i+k) \cos mri + p_m q_m \sin mr(2i+k) + q_m^2 \sin mri \sin mr(i+k).$$

Mas

$$\sum_{i=0}^{N-1} x_i \cos mr(i+k) = \sum_{i=0}^{N-1} x_{i+k} \cos mr(i+2k)$$

donde

$$\sum_{i=0}^{N-1} [x_i \cos mr(i+k) + x_{i+k} \cos mri] =$$

$$= \sum_{i=0}^{N-1} x_i x_{i+k} [\cos mr(i+2k) + \cos mri] =$$

$$= 2 \sum_{i=0}^{N-1} x_i x_{i+k} \cos mr(i+k) \cos mrk =$$

$$= 2 \cos mrk \sum_{i=0}^{N-1} x_i \cos mri = N p_m \cos mrk.$$

Semelhantemente, pode-se verificar que

$$\sum_{i=0}^{N-1} [x_i \sin mr(i+k) + x_{i+k} \sin mri] = N q_m \cos mrk.$$

Tem-se, ainda:

$$\sum_{i=0}^{N-1} \cos mri \cos mr(i+k) = \frac{N}{2} \cos mrk$$

$$\sum_{i=0}^{N-1} \sin mr(2i+k) = 0 \quad (m < N/2)$$

$$\sum_{i=0}^{N-1} \sin mri \sin mr(i+k) = \frac{N}{2} \cos mrk.$$

Resulta

$$S x'_i x'_{i+k} = S x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk, \quad m < N/2.$$

Em resumo, se $m < N/2$,

$$S x'_i x'_{i+k} = S x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk$$

$$S x_i'^2 = S x_i^2 - \frac{N}{2} a_m^2,$$

donde

$$r'_k = \frac{S x'_i x'_{i+k}}{S x_i'^2} = \frac{S x_i x_{i+k} - \frac{N}{2} a_m^2 \cos mrk}{S x_i^2 - \frac{N}{2} a_m^2}.$$

Se N é par e $m = N/2$, não é difícil verificar que

$$r'_k = \frac{S x_i x_{i+k} - N a_m^2 \cos mrk}{S x_i^2 - \frac{N}{2} a_m^2}.$$

Generalização do conceito de coeficiente de correlação cíclica

16. Os coeficientes de correlação cíclica definidos no texto do presente trabalho podem denominar-se mais propriamente, coeficientes de auto-correlação cíclica.

O conceito pode ser generalizado, definindo-se os coeficientes de correlação entre duas sucessões cíclicas,

$$x_0, x_1, x_2, \dots, x_{N-1}$$

$$y_0, y_1, y_2, \dots, y_{N-1},$$

pelas fórmulas

$$r_{xy,k} = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x})(y_{i+k} - \bar{y})}{\sqrt{\sum_{i=0}^{N-1} (x_i - \bar{x})^2 \sum_{i=0}^{N-1} (y_i - \bar{y})^2}}$$

Tem-se, evidentemente,

$$r_{xy,k} = r_{yx,N-k}.$$

Se $y_i \equiv x_i$, obtém-se o coeficiente de auto-correlação cíclica de ordem k .

Um caso particular que encontra aplicação na análise harmônica é o seguinte: dada a sucessão

$$x_0, x_1, x_2, \dots, x_{N-2}, x_{N-1},$$

forme-se a sucessão

$$y_0, y_1, y_2, \dots, y_{N-2}, y_{N-1},$$

onde

$$y_i = x_{N-i}; \\ x_0, x_{N-1}, x_{N-2}, \dots, x_2, x_1.$$

Esta sucessão obtém-se da primeira por inversão e permutação cíclica. Note-se que a soma dos índices dos termos correspondentes nas duas sucessões, x_i e $y_i = x_{N-i}$ é igual a N (ou a zero).

Tem-se

$$y_i x_{i+k} = x_{N-i} x_{i+k}$$

ou, fazendo

$$N - i = j \text{ e, portanto, } i = N - j,$$

$$y_i x_{i+k} = x_j x_{N+k-j} = x_j x_{k-j},$$

de modo que

$$r_{yxk} = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x})(x_{k-i} - \bar{x})}{\sum_{i=0}^{N-1} (x_i - \bar{x})^2}.$$

Determinação da amplitude e do ângulo fase sem recurso aos dados originais

17. Os coeficientes r_{yxk} definidos no parágrafo precedente, em conjunto com os coeficientes de auto-correlação cíclica, $r_k = r_{xx,k}$ permitem determinar p_m^2 e q_m^2 e, portanto, as amplitudes e os ângulos fase das componentes harmônicas, sem necessidade de voltar aos dados originais.²⁹

Tem-se

$$p_m^2 - q_m^2 = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 \cos 2mri + 2 \sum_{i=0}^{N-1} x_i x_j \cos mr(i+j) \right],$$

$$i < j,$$

x_i podendo representar, indiferentemente, os dados originais ou seus desvios em relação à média aritmética.

Reunindo os termos nos quais $i + j = k$ ou $i + j = N + k$, tem-se

$$p_m^2 - q_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} \left(\sum_{i=0}^{N-1} x_i x_{k-i} \right) \cos mrk,$$

donde

$$p_m^2 - q_m^2 = \frac{4}{N^2} \sum_{i=0}^{N-1} (x_i - \bar{x}) \sum_{k=0}^{N-1} r_{yx,k} \cos mrk,$$

onde

$$r_{yx,k} = \frac{\sum_{i=0}^{N-1} (x_i - \bar{x})(x_{k-i} - \bar{x})}{\sum_{i=0}^{N-1} (x_i - \bar{x})^2},$$

como no parágrafo precedente.

²⁹ Infelizmente, os sinais de p_m e q_m não podem ser determinados sem recurso aos dados originais.

Lembrando que

$$p_m^2 + q_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} \left(\sum_{i=0}^{N-1} x_i x_{i+k} \right) \cos mrk,$$

tem-se

$$p_m^2 + q_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} R_{xx,k} \cos mrk,$$

$$p_m^2 - q_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} R_{yx,k} \cos mrk,$$

onde

$$R_{xx,k} = \sum_{i=0}^{N-1} x_i x_{i+k}$$

e

$$R_{yx,k} = \sum_{i=0}^{N-1} x_i x_{k-i}.$$

Obtém-se

$$N^2 a^2/4 = \sum_{k=0}^{N-1} R_{xx,k} \cos mrk,$$

$$2 p_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} (R_{xx,k} + R_{yx,k}) \cos mrk,$$

$$2 q_m^2 = \frac{4}{N^2} \sum_{k=0}^{N-1} (R_{xx,k} - R_{yx,k}) \cos mrk,$$

donde

$$p_m^2/q_m^2 = \frac{\sum_{k=0}^{N-1} (R_{xx,k} + R_{yx,k}) \cos mrk}{\sum_{k=0}^{N-1} (R_{xx,k} - R_{yx,k}) \cos mrk},$$

permitindo calcular as amplitudes e os ângulos fase com grande exatidão, sem voltar aos dados originais.³⁰

Na prática, é, portanto, desnecessário calcular os coeficientes de auto-correlação cíclica $r_{xx,k}$ e os coeficientes de correlação cíclica $r_{yx,k}$. É suficiente calcular, usando o mesmo dispositivo mecânico descrito no texto,

$$R_{xx,k} = \sum_{i=0}^{N-1} x_i x_{i+k}$$

$$R_{yx,k} = \sum_{i=0}^{N-1} x_i x_{k-i},$$

onde os x_i representam os dados originais.

Tem-se

$$R_{xx,N-k} = R_{xx,k} \\ \sum_{k=0}^{N-1} R_{xx,k} = \left(\sum_{i=0}^{N-1} x_i \right)^2,$$

³⁰ Infelizmente, os sinais de p_m e q_m não podem ser determinados sem recurso aos dados originais.

relações que permitem abreviar e conferir a correção dos cálculos. Verifica-se, também,

$$\sum_{k=0}^{N-1} R_{yx,k} = \left(\sum_{i=0}^{N-1} x_i \right)^2$$

A harmônica predominante é indicada pelo mínimo absoluto de $R_{xx,k}$, dispensando a construção do correlograma.

A eliminação da onda predominante, de frequência m/N , faz-se pelas fórmulas

$$R_{xx,k} = R_{xx,k} - \frac{N}{2} a_m^2 \cos mrk,$$

$$R'_{yx,k} = R_{yx,k} - k^{21} \frac{N}{2} (p_m^2 - q_m^2) \cos mrk - \\ - N p_m q_m \sin mrk^{21}$$

Componentes q das harmônicas de $r_{yx,k}$

²¹ Deixa-se ao leitor a demonstração desta última fórmula.

18. Como foi visto, as componentes q das harmônicas dos coeficientes de auto-correlação cíclica são tôdas nulas. O mesmo não acontece com os coeficientes $r_{yx,k}$ definidos no item 16.

Tem-se, efetivamente,

$$p_m q_m = \frac{4}{N^2} \left[\sum_{i=0}^{N-1} x_i^2 \sin 2 mri + \right. \\ \left. + 2 \sum_{i=0}^{N-1} x_i x_{i+j} \sin mr(i+j) \right], i < j.$$

Reunindo os termos com $i + j = k$ ou $i + j = N + k$, vem

$$p_m q_m = \frac{4}{N^2} \sum_{k=0}^{N-1} \left(\sum_{i=0}^{N-1} x_i x_{k-i} \right) \sin mrk,$$

de modo que a componente q de $r_{yx,k}$ é nula se, e somente se, $p_m = 0$ ou $q_m = 0$, como acontece quando N é par e $m = N/2$.