On Periodicity in Series of Related Terms.

By Sir Gilbert Walker, F.R.S.

(Received January 22, 1931.)

An important extension of our ideas regarding periodicity was made in 1927 when Yule* pointed out that, instead of regarding a series of annual sunspot numbers as consisting merely of a harmonic series to which a series of random terms were added, we might suppose a certain amount of causal relationship between the successive annual numbers. In that case the system might be regarded as a physical system possessing one or more natural oscillations of its own, all subject to damping; and the effect of annual random disturbances would be to produce a fairly smooth curve with periods varying in amplitude and length, essentially as the sunspot numbers vary. If we call the departures from their mean of our series u_1, u_2, ..., Yule showed that the consequence of a single natural period is an equation like

$$u_x = ku_{x-1} - u_{x-2} + v_x,$$

where v_x represents the "accidental" external "disturbance"; and if there are two natural periods,

$$u_x = k_1 (u_{x-1} + u_{x-3}) - k_2 u_{x-2} - u_{x-4} + v_x.$$

He considered also the effect of the relation

$$u_x = g_1 u_{x-1} - g_2 u_{x-2}, \tag{A}$$

which leads to a damped harmonic vibration

$$u_x = e^{-\lambda x} (\mathrm{A} \cos \theta x + \mathrm{B} \sin \theta x),$$

where exp. $(-\lambda \pm i\theta)$ are the roots of the equation

$$y^2 - g_1 y + g_2 = 0.$$

Yule determined his constants by applying the equation (A) to the successive terms of the u series and using the method of least squares.

2. We shall now consider such serial correlation coefficients as r_1, r_2, ..., where r_1 is that between consecutive terms of u's and r_p that between terms

* 'Phil. Trans.,' A, vol. 226, pp. 267–298 (1927).

separated by p intervals. The use of such coefficients in connection with periodicities is old, but it has recently been more widely employed.*

If the equation connecting successive terms in the absence of disturbance is

$$u_x = g_1 u_{x-1} + g_2 u_{x-2} + \dots + g_s u_{x-s}, \tag{B}$$

then the equation for the terms as disturbed is obtained by adding a term v_x to the right-hand side. Let us multiply this equation by u_{x-s-1} and sum for all values of x from $(s + 2)$ to n ; we get, ignoring the sums of product terms in uv as relatively insignificant because the v's are accidental,

$$\sum_{s+2}^{n} \{u_x u_{x-s-1} - (g_1 u_{x-1} u_{x-s-1} + \dots + g_s u_{x-s} u_{x-s-1})\} = 0. \tag{C}$$

In order to simplify the analysis we treat the number of terms as so large that we may neglect errors due to its finiteness ;† then if the S.D. of the series is d, we can neglect the differences between the S.D.'s of n terms and of $(n - s)$ terms, so that (C) becomes

$$(n - s - 1)\, d^2 \{r_{s+1} - (g_1 r_s + g_2 r_{s-1} + \dots + g_s r_1)\} = 0,$$

or

$$r_{s+1} = g_1 r_s + g_2 r_{s-1} + \dots + g_s r_1.$$

Similarly on multiplying by u_{x-s-2} and adding,

$$r_{s+2} = g_1 r_{s+1} + g_2 r_s + \dots + g_s r_2 ;$$

and in general

$$r_y = g_1 u_{y-1} + g_2 u_{y-2} + \dots + g_s r_{y-s},$$

which is analogous with (B) and shows that the relationships between the successive u's and the successive r's are, in the limit when n is very large, identical.

If the roots of the equation

$$z^s = g_1 z^{s-1} + \dots + g_s \tag{D}$$

are h_1, h_2 ... h_s the solution of (B) will be

$$u_p = \mathrm{U}_1 h_1^{\,p} + \mathrm{U}_2 h_2^{\,p} + \dots + \mathrm{U}_s h_s^{\,p}$$

* See, for example, Dinsmore Alter in the Washington 'Monthly Weather Review,' June, 1927 ; and a derivation of a criterion of reality applicable to H. H. Turner's "chapters" of continuous oscillations, separated by breaks due to outside interference, in paragraph 11, pp. 340, 341 of a paper "On Periodicity," 'Q.J.R. Met. Soc.,' vol. 51 (1925).

† For instance, it is customary, if seeking a 13-year period with 120 annual values available, to consider 117 years out of the 120, and to assume that ignoring 3 terms will not seriously affect the result.

where the U's are constants; and of (D) will be

$$r_p = R_1 h_1^p + R_2 h_2^p + \dots + R_s h_s^p$$

where the R's are constants.

3. Thus the r's must have the same periods as the u's. This is obvious in slightly damped simple oscillations each occupying say q of the intervals between the u's; for then $r_{q+1}, r_{q+2} \dots r_{2q}$ will tend to be the same as $r_1, r_2, \dots r_q$; the r's will thus have a period of q intervals and will be damped if the u's are damped. One advantage of using the values of $r_1, r_2 \dots$ for getting the relationship (B) over using the u's is that the former, being based on the whole series, are much less influenced by accidental effects.

4. We shall first of all consider in somewhat greater detail the effect of the simplest type of dependence of each term on the previous ones, that of the tendency of a departure to persist; and we shall suppose that any term u_s is made up of tu_{s-1} (where t is a fraction less than unity) and of an external 'disturbance' v_s. Also, since we may write $u_s = tu_{s-1} + v_s$ in the form $u_s - u_{s-1} = -(1-t)\, u_{s-1} + v_s$, we might regard "persistence" as equivalent to "damping" in a mechanical system, the diminution being proportional to the magnitude of the previous term. If now we write the equations

$$u_2 = tu_1 + v_2, \quad u_3 = tu_2 + v_3, \quad \dots \; u_n = tu_{n-1} + v_n,$$

and assume that n is so large that the S.D.'s d and d' of the u's and v's may be treated as unaffected by cutting out a term, we realise (1) that t is the correlation coefficient that we have previously denoted by r_1, and (2) that by the ordinary theorem

$$d'^2 = (1 - r_1^2)\, d^2. \tag{E}$$

Now for the Fourier terms of the u series

$$\frac{n}{2}(a_q + ib_q) = \sum_{k=1}^{n} u_k e^{i(k-1)q\alpha},$$

where $\alpha = 2\pi/n$; and for the Fourier terms of the series of disturbances, which terms are distinguished by dashes, inasmuch as u_0 is unknown, v_1 is indeterminate; but when n is large enough we may make any hypothesis we like regarding v_1 without appreciably affecting the Fourier terms; and we choose it as equal to $u_1 - r_1 u_n$. Thus

$$\frac{n}{2}(a_q' + ib_q') = \sum_{k=1}^{n} v_k e^{i(k-1)q\alpha},$$

and, on substituting $u_k - r_1 u_{k-1}$ for v_k, it is easily seen that this becomes

$$(1 - r_1 e^{iq\alpha}) \sum_{l=1}^{n} u_l e^{i(l-1)q\alpha}, \quad \text{or} \quad \frac{n}{2}(1 - r_1 e^{iq\alpha})(a_q + ib_q).$$

So multiplying by the corresponding equation with the sign of i changed, if $c'^2 = a'^2 + b'^2$ and $c^2 = a^2 + b^2$,

$$c_q'^2 = c_q^2 (1 - 2r_1 \cos q\alpha + r_1^2). \qquad \text{(F)}$$

Also if $c'/2^{\frac{1}{2}}d'$, the amplitude ratio,* is denoted by f' and $c/2^{\frac{1}{2}}d$ by f, we have using (E)

$$f_q^2/f_q'^2 = (1 - r_1^2)/(1 - 2r_1 \cos q\alpha + r_1^2). \qquad \text{(G)}$$

It would appear that if we had two physical systems, one in which the successive values were independent and a second system in which persistence produced a relationship r between successive terms, and if the same disturbances were imposed on the systems, then, by (F), when the oscillations had gone on so long that a fairly steady mean amplitude had been attained in the persistent system, the amplitude c' of the first or free system would average $(1 - 2r \cos q\alpha + r^2)^{\frac{1}{2}}$ times that of the second or persistent system, and the persistence would alter the amplitude ratio of the oscillations set up in the ratio $(1 - r^2)^{\frac{1}{2}}/(1 - 2r \cos q\alpha + r^2)^{\frac{1}{2}}$.

It may be noted that q lies between 1 and $n/2$, and $q\alpha$ between $2\pi/n$ and π, and so between 0 and π. Thus the ratio f^2/f_1^2 lies between $(1 + r)/(1 - r)$ and $(1 - r)/(1 + r)$; it is unity when $\cos q\alpha = r$. In practice q does not in general exceed $n/6$, so $q\alpha$ is not in general greater than $\pi/3$. For instance, with 120 annual values, for a period of 20 years $q\alpha = \pi/10$, and for 8 years it is $\pi/4$.

The result that "persistence" or inertia will diminish the amplitudes of oscillations of short period, but may increase the relative importance, and therefore the amplitude ratio, of those of long period, by destroying quick oscillations, is in accordance with expectation.

5. Let us now consider the oscillations set up in a system which has natural periods of its own. Corresponding to (D), with $s = 3$ for brevity, we have the equations

$$\left.\begin{aligned} u_4 &= g_1 u_3 + g_2 u_2 + g_3 u_1 + v_4 \\ u_5 &= g_1 u_4 + g_2 u_3 + g_3 u_2 + v_5 \\ u_n &= g_1 u_{n-1} + g_2 u_{n-2} + g_3 u_{n-3} + v_n \end{aligned}\right\}, \qquad \text{(H)}$$

* The amplitude ratio of a Fourier term is the ratio of its amplitude to $2^{\frac{1}{2}}$ times the standard deviation of the terms analysed, see (*c*) p. 26 of 'Q.J.R. Met. Soc.,' vol. 54 (1928). If a series consists accurately of a single sine series the ratio is unity.

and we may define quantities v_1, v_2, v_3 by the equations

$$\left.\begin{aligned} u_1 &= g_1 u_n + g_2 u_{n-1} + g_3 u_{n-2} + v_1 \\ u_2 &= g_1 u_1 + g_2 u_n + g_3 u_{n-1} + v_2 \\ u_3 &= g_1 u_2 + g_2 u_1 + g_3 u_n + v_3 \end{aligned}\right\}, \qquad \text{(I)}$$

which would be continuous with (H) if the u series repeated itself after n terms, so that $u_{n+1} = u_1$, $u_{n+2} = u_2$, etc. Then the typical Fourier terms of the u and v series will be

$$a_q \cos qx + b_q \sin qx \quad \text{and} \quad a_q{}' \cos qx + b_q{}' \sin qx$$

where

$$\begin{aligned} a_q + ib_q &= \frac{2}{n} \sum_{p=0}^{n-1} u_{p+1}\, e^{ipq\alpha} \\ &= \frac{2}{n} \sum_{0}^{n-1} \{g_1 u_p\, e^{ipq\alpha} + g_2 u_{p-1}\, e^{ipq\alpha} + g_3 u_{p-2}\, e^{ipq\alpha} + v_{p+1}\, e^{ipq\alpha}\} \\ &= (a_q + ib_q)\{g_1\, e^{iq\alpha} + g_2\, e^{2iq\alpha} + g_3\, e^{3iq\alpha}\} + a'_q + ib'_q. \end{aligned}$$

Therefore

$$a_q + ib_q = (a'_q + ib'_q)/(1 - g_1\, e^{iq\alpha} - g_2\, e^{2iq\alpha} - g_3\, e^{3iq\alpha})$$

so

$$c_q{}^2 = c'_q{}^2/\{1 + g_1{}^2 + g_2{}^2 + g_3{}^2 + 2\,(g_1 g_2 + g_2 g_3 - g_1)\cos\alpha$$
$$+ 2(g_1 g_3 - g_2)\cos 2\alpha - 2g_3 \cos 3\alpha\}. \qquad \text{(J)}$$

6. Further it may be seen that if there is a natural period of the u's corresponding to this Fourier term in the accidental disturbances and the damping is small, the amplitude set up will be relatively large. For if there is a natural period of the u's, in view of which they repeat after n/q terms, we have to consider n/q in conjunction with the solution exp. $(-\lambda \pm i\theta)$ of the equation $x^3 - g_1 x^2 - g_2 x - g_3 = 0$, the undisturbed u terms being got by giving values 1, 2, 3, ..., to p in $e^{-\lambda p}\,(\mathrm{A}\cos p\theta + \mathrm{B}\sin p\theta)$. Thus, if λ is small, $2\pi/\theta$ must be n/q. But as exp. $(-\lambda \pm i\theta)$ satisfied the cubic equation the value of

$$e^{\pm 3i\theta} - g_1\, e^{\pm 2i\theta} - g_2\, e^{\pm i\theta} - g_3$$

will be small, and as $\theta = 2\pi q/n = q\alpha$ the equation preceding (J) shows that c_q will be relatively large; thus the u system will, if its damping is not too large, act like a resonator and respond, in its own periods, to relatively small accidental external disturbances.

7. If we want the ratio of the mean magnitudes of the accidental disturbances v to the u's of the original series, we realise that the equations (H) and (I)

may be interpreted algebraically as a regression equation with coefficients g_1, g_2, g_3 by means of which the terms of the series $u_4, u_5, \ldots u_n, u_1, u_2, u_3$ are expressed as linear functions of the terms in the three series $(u_3, u_4, \ldots, u_2)$, $(u_2, u_3, \ldots u_1)$, and $(u_1, u_2, \ldots u_n)$; so the joint correlation coefficient R, between the first series u_x and the series of which the general term is

$$(g_1 u_{x-1} + g_2 u_{x-2} + g_3 u_{x-3}),$$

is given by*

$$R^2 = g_1 r_1 + g_2 r_2 + g_3 r_3,$$

and then, as v_4 is independent of u_1, u_2, u_3 and v_5 of u_2, u_3, u_4, etc., we have the same algebraic relation as in ordinary statistics,

$$\sum_1^n v_x^2 = (1 - R^2) \sum_1^n u_x^2$$

or

$$d'^2 = (1 - g_1 r_1 - g_2 r_2 - g_3 r_3)\, d^2. \tag{K}$$

8. In a practical application of the method we work out the series of values of r_p and, if their graph clearly contains oscillations with certain periods, it is conceivable that instead of these being all natural periods of the original system some might be due to periodicities in the external disturbances which would no longer be purely accidental. Accordingly we shall examine the case in which the external disturbances indicated by v_p are made up of two portions, one f_p a periodic function of p and the other w_p a purely "accidental" element. We may then express u_p as given by the relation

$$u_p = g_1 u_{p-1} + g_2 u_{p-2} + \ldots + g_s u_{p-s} + f_p + w_p.$$

Let the roots of the equation

$$y^s = g_1 y^{s-1} + \ldots + g_s$$

be $\alpha_1, \alpha_2, \ldots, \alpha_s$ and let the oscillation denoted by f_p be governed by a similar equation

$$y^t = h_1 y^{t-1} + \ldots + h_t$$

with roots $\beta_1, \beta_2, \ldots \beta_t$. Then f_p is of the form $B_1\beta_1^p + B_2\beta_2^p + \ldots + B_t\beta_t^p$; and if the equation whose roots are all the α's and all the β's is

$$y^{s+t} = k_1 y^{s+t-1} + k_2 y^{s+t-2} + \ldots + k_{s+t}$$

we shall have the relations

$$u_p = k_1 u_{p-1} + \ldots + k_{s+t} u_{p-s-t} + w_p$$

and

$$r_p = k_1 r_{p-1} + \ldots + k_{s+t} r_{p-s-t}.$$

* 'Indian Meteorological Memoirs,' vol. 20, p. 122, equation 2 (1908).

Thus when we plot the r_p graph we shall see in it the oscillations of both the internal and external systems. Now it often happens that from the nature of the case the oscillations of the external disturbing system are undamped, while those of the disturbed system must be damped; and then the interpretation of the graph should be possible.

9. Some light is thrown by this analysis on the utility of a series of values of r_p, a "correlation-periodogram," as a substitute for the ordinary Fourier-Schuster periodogram when there is no question of damped oscillations.

We shall first consider the relations between the r's of the former and the f's, the amplitude-ratios, of the latter. If the number of departures $u_1, u_2, \ldots u_n$ be $2m+1$, as usual

$$u_p = a_0 + a_1 \cos p\alpha + \ldots + a_m \cos mp\alpha$$
$$+ b_1 \sin p\alpha + \ldots + b_m \sin mp\alpha,$$

and $a_0 = 0$ since the series consists of departures from the mean. Thus in general when n is large enough for the S.D. d to be unaffected by modifying a few u terms

$$r_s = \sum_{p=1}^{n-s} u_p u_{p+s}/(n-s)d^2.$$

Also as n is supposed large while s remains finite we may as a first approximation replace this expression by

$$r_s = \sum_{p=1}^{n} u_p u_{p+s}/nd^2,$$

it being assumed, as before, that $u_{n+s} \equiv u_s$; thus

$$r_s = \frac{1}{nd^2}\sum_{p=1}^{n}\left\{\begin{array}{l} a_1 \cos p\alpha + \ldots \\ \qquad + a_m \cos mp\alpha \\ + b_1 \sin p\alpha + \ldots \\ \qquad + b_m \sin mp\alpha \end{array}\right\}\left\{\begin{array}{l} a_1 \cos (p+s)\alpha + \ldots \\ \qquad + a_m \cos m(p+s)\alpha \\ + b_1 \sin (p+s)\alpha + \ldots \\ \qquad + b_m \sin m(p+s)\alpha \end{array}\right\}$$

$$= \frac{1}{nd^2}\{(a_1^2+b_1^2)\cos s\alpha + (a_2^2+b_2^2)\cos 2s\alpha + \ldots\}\frac{n}{2}$$

$$= f_1^2 \cos s\alpha + f_2^2 \cos 2s\alpha + \ldots + f_m^2 \cos ms\alpha. \tag{L}$$

A partial check is easy; for if the series forms an accurate cosine curve with a period of n/q or s terms, let us say, where s is an integer, the property of amplitude-ratios tells us that $f_q = 1$, all the other f's vanishing; and as the series repeats itself completely after s terms we shall have $r_s = 1$.

Thus any period of q terms with an amplitude ratio f will produce as graph for r_p a cosine curve with maxima of at f_p at $p = q, 2q, 3q, \ldots$, and equal and opposite minima half-way between.

Accordingly if there are only one or two periods and they are well-marked, inspection of the correlation-periodogram will reveal them; but if there are three or four periods or they are ill-marked, Fourier analysis of the r_p curve will be necessary.

10. We will now apply these ideas to the pressure at Port Darwin, one of the most important centres of action of "world weather," which, like the closely related station of Batavia, displays surges of varying amplitude and period with irregularities superposed, suggesting that pressure in this region has a natural period of its own, based presumably on the physical relationships of world-weather, but that the oscillations are modified by external disturbances. The data examined have been the 177 quarterly pressure values from 1882 to 1926, and as a first experiment we have considered the amplitude ratio f of the 26 Fourier harmonics* from the 5th to the 30th, covering periods from 9 years to $1\frac{1}{2}$ years, which will be found in Table I; the pressure curve and the

Table I.—Periodogram of Port Darwin Pressure 177 quarters.

Order of harmonic.	Fourier period.	Period examined.	Amplitude ratio.	Amplitude ratio corrected.
q.	Quarters.	Quarters.	f.	f'.
5	35·4	35	0·19	0·08
6	29·5	30	0·06	0·03
7	25·3	25	0·19	0·09
8	22·1	22	0·24	0·13
9	19·7	20	0·14	0·08
10	17·7	17·5	0·21	0·13
11	16·1	16	0·12	0·08
12	14·7	14·7	0·18	0·12
13	13·6	13·5	0·29	0·21
14	12·6	12·5	0·07	0·05
15	11·8	11·7	0·29	0·23
16	11·1	11	0·24	0·20
17	10·4	10·5	0·15	0·13
18	9·8	10	0·08	0·07
19	9·3	9·3	0·14	0·13
20	8·8	8·8	0·06	0·06
21	8·4	8·3	0·02	0·02
22	8·0	8	0·11	0·12
23	7·69	7·66	0·14	0·16
24	7·38	7·33	0·07	0·08
25	7·08	7	0·10	0·12
26	6·80	6·67	0·04	0·05
27	6·56	6·5	0·03	0·04
28	6·32	6·29	0·05	0·07
29	6·10	6·12	0·16	0·22
30	5·90	5·90	0·07	0·10

* The periods examined are sufficiently close to the Fourier periods; see pp. 119, 120, of 'Memoirs of R. Met. Soc.,' I, No. 9, 1927.

periodogram are in figs. 1 and 2. It will be seen that the 8th, 13th, 15th and 16th harmonics have the largest ratios ; their periods are 22, $13\frac{1}{2}$, $11\frac{2}{3}$ and 11

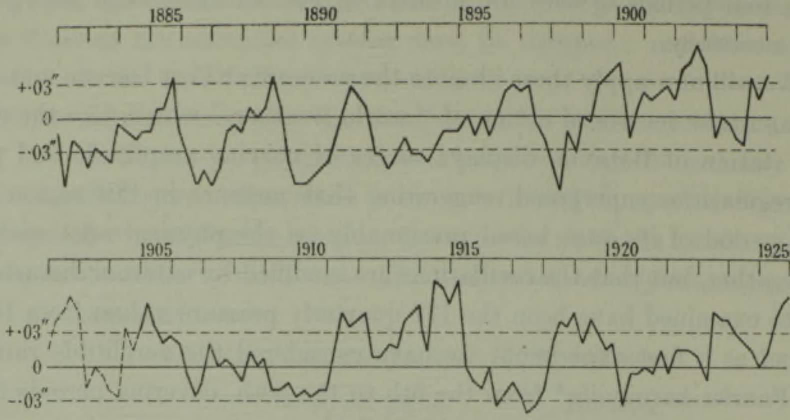

Fig. 1.—Port Darwin Pressure.

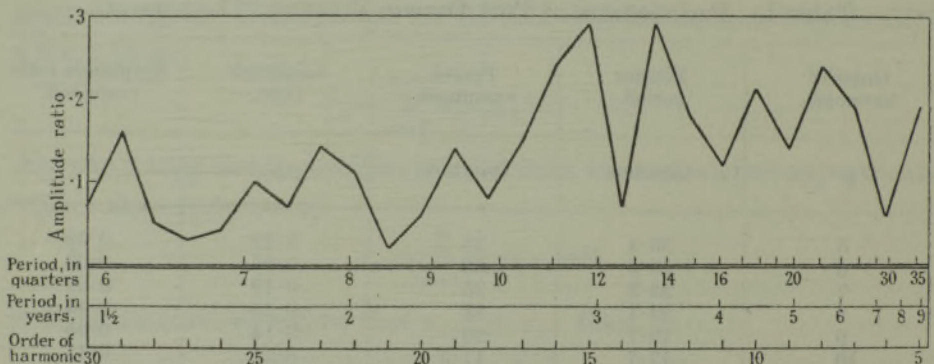

Fig. 2.—Periodogram.

quarters, and their f's are 0·24, 0·29, 0·29 and 0·24, the amplitude ratio for a pure sine curve being unity. The probable value of a single f is 0·088, and the probable value* of the greatest of 26 of these, if they had been independent, would have been 0·20. But inasmuch as the correlation coefficient r_1 between successive quarterly pressures is 0·76, the terms of the series to be analysed are far from independent ; and we cannot compare the results directly with those derived by Fourier analysis of a random series. It may be noted that the amplitudes of the 15th and 16th harmonics point to a single natural period of intermediate length, say, $11\frac{1}{3}$ quarters or 2·8 years, while the length of the 13th harmonic is 3·4 years.

11. But granted the persistence we naturally interpret the pressure variations

* 'Q.J.R. Met. Soc.,' para. 6, p. 338, vol. 51 (1928).

in one of two ways. Either (*a*) the pressure is like a mechanical system, with persistence but without natural periods and acted on by a series of disturbances; in this case it is the periodicity of the disturbances that must be examined. Or (*b*) the pressure behaves like a mechanical system with persistence and natural periods, and then these periods interest us. In the first case having found $r = 0 \cdot 76$ for Port Darwin we may* use equation (G) above and deduce the amplitude ratio f' of the disturbances from the ratios f of the original series. These are given in the last column of Table I† and plotted in fig. 3; we have four ratios, corresponding to the 13th, 15th, 16th and 29th

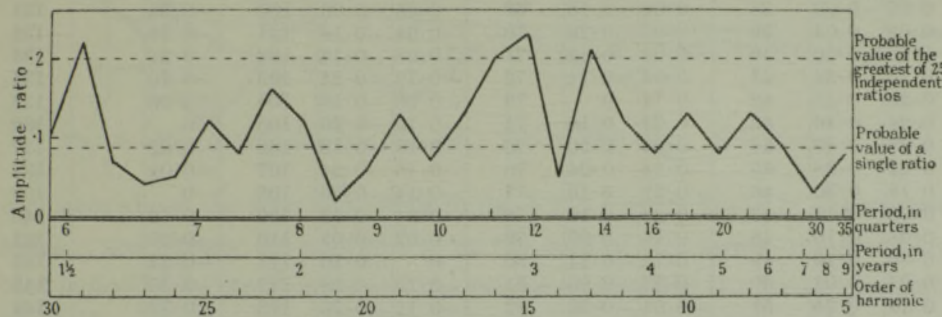

FIG. 3.—Periodogram of Disturbances, when Persistence alone is allowed for.

harmonics, which reach the limit of $0 \cdot 20$ that we should expect as the largest produced by mere chance. The ratio of the 14th harmonic is only $0 \cdot 07$. Regarding the amplitudes of the 15th and 16th harmonics as due to a single period of about $11\frac{1}{3}$ quarters, this interpretation suggests that while belief in the reality of three periods, of about $3 \cdot 4$ years, $2 \cdot 8$ years and $1 \cdot 53$ years, is permissible, it is far from inevitable.

12. We shall now consider the second interpretation and examine the idea that Port Darwin pressure has natural periods of its own, maintained by non-periodic disturbances from outside. The values of the correlation coefficients between quarterly pressures separated by an interval of p quarters are given in Table II from $p = 1$ to $p = 147$; but, before discussing these in full, it may be well to gain experience of the method in a preliminary trial of a simpler problem, and consider the first 40 coefficients which are plotted in fig. 4 in the continuous curve A. In this first experiment, then, we have to determine r_p

* It will be seen that in paragraph 4 above it has not been assumed that the disturbances are accidental.

† As a check the values of the successive v's have been tabulated, and 10 of their amplitude ratios calculated; these all agreed with those derived as above within $0 \cdot 03$, *i.e.*, within the limit of probable error.

Table II.—Correlation Coefficients of Port Darwin Pressure one Quarter with another after various intervals, (*a*) using all data, (*b*) using only 77 pairs.

Interval.	All data.	77 pairs.	Interval.	All data.	77 pairs.	Interval.	All data.	77 pairs.	Interval.	All data.	77 pairs.	Interval.	All data.
0	1·00	1·00	31	0·12	0·22	62	—0·22	—0·20	93	0·22	0·12	124	—0·30
1 *qr.*	0·76	0·80	32	0·06	0·22	63	—0·22	—0·26	94	0·12	0·04	125	—0·44
2	0·56	0·58	33	0·10	0·24	64	—0·14	—0·18	95	0·08	0	126	—0·62
3	0·36	0·34	34	0·10	0·16	65	—0·10	—0·20	96	—0·04	—0·10	127	—0·68
4	0·18	0·12	35	0·12	0·08	66	—0·08	—0·22	97	—0·16	—0·22	128	—0·60
5	0·08	0·02	36	0·08	—0·06	67	0·02	—0·10	98	—0·16	—0·20	129	—0·58
6	0·02	—0·02	37	0·06	—0·16	68	0·06	—0·04	99	—0·30	—0·32	130	—0·48
7	0·02	—0·02	38	0·06	—0·18	69	0·06	—0·06	100	—0·34		131	—0·38
8	—0·02	0·04	39	0·02	—0·22	70	—0·04	—0·14	101	—0·34		132	—0·38
9	0·08	0·20	40	0·08	—0·18	71	—0·08	—0·18	102	—0·24		133	—0·24
10	0·16	0·34	41	0·04	—0·14	72	—0·12	—0·24	103	—0·10		134	—0·12
11	0·22	0·48	42	0·14	0	73	—0·06	—0·16	104	—0·06		135	—0·04
12	0·24	0·46	43	0·24	0·10	74	—0·10	—0·20	105	0		136	0
13	0·28	0·48	44	0·26	0·20	75	—0·10	—0·18	106	—0·02		137	0·08
14	0·22	0·38	45	0·28	0·24	76	—0·16	—0·24	107	—0·04		138	0·06
15	0·18	0·30	46	0·22	0·16	77	—0·14	—0·20	108	0		139	—0·14
16	0·08	0·18	47	0·20	0·10	78	—0·04	—0·12	109	—0·10		140	—0·22
17	0·08	0·16	48	0·16	0·02	79	—0·02	—0·08	110	—0·20		141	—0·36
18	0·06	0·10	49	0·12	—0·14	80	0	—0·10	111	—0·24		142	—0·44
19	0·04	0·08	50	0·02	—0·20	81	—0·10	—0·20	112	—0·30		143	—0·56
20	0·14	0·24	51	—0·04	—0·22	82	—0·12	—0·26	113	—0·32		144	—0·60
21	0·12	0·30	52	—0·08	—0·26	83	—0·20	—0·40	114	—0·40		145	—0·60
22	0·18	0·34	53	—0·02	—0·16	84	—0·20	—0·38	115	—0·36		146	—0·48
23	0·20	0·36	54	0·04	—0·10	85	—0·22	—0·40	116	—0·34		147	—0·36
24	0·20	0·28	55	0·12	0·04	86	—0·24	—0·36	117	—0·30			
25	0·18	0·14	56	0·08	0·02	87	—0·20	—0·30	118	—0·22			
26	0·10	0	57	0·04	0·04	88	—0·14	—0·22	119	—0·28			
27	0·06	—0·12	58	0	0·06	89	—0·10	—0·18	120	—0·30			
28	0·04	—0·10	59	—0·14	—0·08	90	0	—0·06	121	—0·22			
29	0·10	0	60	—0·16	—0·08	91	0·12	0·04	122	—0·16			
30	0·12	0·12	61	—0·22	—0·16	92	0·16	0·04	123	—0·22			

as the sum of such terms as $e^{-\alpha p} \cos(\beta p + \gamma)$ in such a way as to give a fair approximation to the curve.

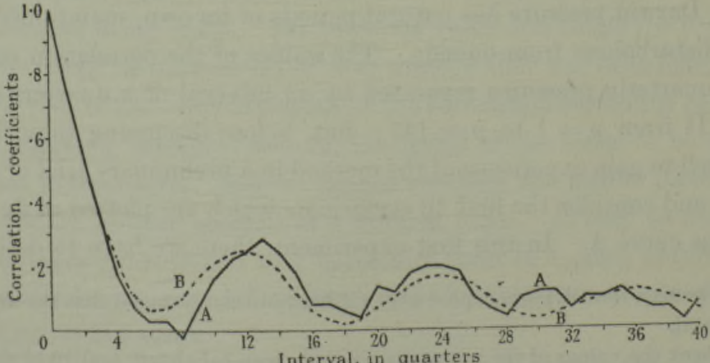

Fig. 4.—Preliminary Experiment. Serial Correlation Coefficients of Port Darwin Pressure, ——, A, actual values of r_1, r_2 ,.... - - - -, B, sum of the ordinates P, Q, R, of fig 5.

Now inspection shows a rapid descent from unity at $p = 0$ followed by maxima near $p = 13$ and 23 ; the natural interpretation is in terms of a damped harmonic curve with a period of about 12 quarters, producing maxima of diminishing amplitudes for $p = 12$ and 24 ; so we take $r_p = 0 \cdot 19\ (0 \cdot 96)^p \cos 2p/12$, plotted as curve P in fig. 5, as a first approximation. But this

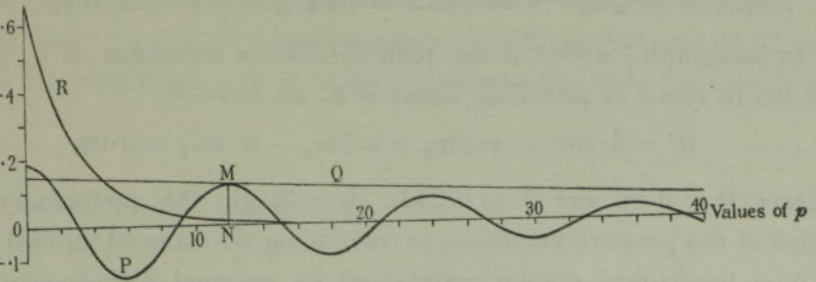

FIG. 5.—$P = 0 \cdot 19(0 \cdot 96)^p \cos (2\pi p/12)$. $Q = 0 \cdot 15(0 \cdot 98)^p$. $R = 0 \cdot 66(0 \cdot 71)^p$.

oscillates about the zero line, while in fig. 4 the **r** curve has only one negative value ; accordingly we interpret r_p as the sum of two terms, the first being that indicated by the P curve and the second a curve like Q whose co-ordinates diminish gradually ; that plotted as Q is $r_p = 0 \cdot 15 \times (0 \cdot 98)^p$, which we may regard as a slightly damped oscillation of infinite period. Corresponding to this there must be a similar term in the expression for u_p, the typical term in the pressure curve, and a sloping line there would mean what is sometimes called a " secular change " in the pressure. The existence of such a change is obvious from fig. 1 and the correlation coefficient $0 \cdot 43$ of pressure with time is one of the largest in the world.*

The sum of these two terms gives fair agreement except that from $p = 0$ to $p = 6$ the curve has not large enough co-ordinates and does not descend fast enough. So we add a strongly damped curve R of which the ordinate is $0 \cdot 66\ (0 \cdot 71)^p$, the effect of which is to increase the amount of " persistence " in the pressure ; the coefficient $0 \cdot 66$ is chosen so as to make $r_0 = 1$. The sum of P, Q, R is indicated by the dotted graph B in fig. 4, which is a fair approximation. Thus we find

$$r_p = 0 \cdot 19\ (0 \cdot 96)^p \cos 2\pi p/12 + 0 \cdot 15\ (0 \cdot 98)^p + 0 \cdot 66\ (0 \cdot 71)^p, \qquad \text{(M)}$$

and the equation corresponding to equation (D) in paragraph 2 above is

$$z^p = 3 \cdot 35 z^{p-1} - 4 \cdot 43 z^{p-2} + 2 \cdot 71 z^{p-3} - 0 \cdot 64 z^{p-4}. \qquad \text{(N)}$$

* It may be that this change is partly due to some change of barometric correction, but its effect on periodicity will be insignificant.

Apart from external disturbances the corresponding solution for the u series would be of the form

$$u_p = (0\cdot96)^p (\mathrm{A} \cos 2\pi p/12 + \mathrm{B} \sin 2\pi p/12) + \mathrm{C}\,(0\cdot98)^p + \mathrm{D}\,(0\cdot71)^p$$

and the effect of disturbances v_1, v_2, .., is to give a typical equation

$$u_p = 3\cdot35u_{p-1} - 4\cdot43u_{p-2} + 2\cdot71u_{p-3} - 0\cdot64u_{p-4} + v_p \tag{O}$$

so, as in paragraph 7 above, if the joint correlation coefficient of u_p with its expression in terms of preceding terms is R, we have

$$\mathrm{R}^2 = 3\cdot35r_1 - 4\cdot43r_2 + 2\cdot71r_3 - 0\cdot64r_4 = 0\cdot92,$$

and hence $\mathrm{R} = 0\cdot96$ and $d' = 0\cdot28d$. Accordingly this preliminary interpretation of the pressure variations as resembling the damped natural oscillations of a mechanical system maintained by external disturbances would explain a very large fraction of the variations, the magnitude of the disturbances averaging only about a quarter of that of the oscillations. As we shall see, however, further extension leads to a different result, and shows the danger of an incomplete examination.

13. Regarding the periods of the natural oscillations it will be seen that the curve A of fig. 4 is not capable of resolving the difference between the damped oscillations of 2·8 and 3·4 years. A Fourier analysis of it would have a third harmonic with a period of $3\frac{1}{3}$ years and a fourth of $2\frac{1}{2}$ years ; the intermediate amplitudes would, as Turner showed, not be independent of these. On this account and in order to discuss the period of between 11 and 12 years suggested by the values of 4 near $p = 44$, 93 and 137 we must extend our examination of the values of r_p, as in curve A, fig. 6 ; here when p is 20 we have 157 pairs of correlates, but as p grows the number of correlates diminishes until when $p = 140$ it is only 30. A glance shows outstanding oscillations near $p = 44$, 93 and 137 with three smaller oscillations between 0 and 44, three between 44 and 93, and two between 93 and 137 ; the general downward slope is maintained. The obvious interpretation is that we have an oscillation with eight periods in 92 quarters, fitting well with the intermediate maxima ; superposed on this there is evidence of a rise up to maxima near 46, 92 and 137 with minima in between, or of an oscillation with a period of about 46 quarters. But far from showing damping the oscillations grow with p, and the explanation seems to lie in the contrast between the number $(177 - p)$ of correlates when p is small and the number when p exceeds 100. Thus for the last 40 terms the number of values correlated averages 50, covering $12\frac{1}{2}$ years, and we have the first 12 or 13 years correlated, with different lags, with the last 12 or 13 years ;

as fig. 1 shows, each has well-marked waves and it is obvious that there will be relative positions in which the waves correspond; so there will be big oscillations in r_p on a scale that would not arise if the number of years correlated were longer. In order to verify this the columns in Table II headed " 77 pairs " have been computed, giving the values of r_p derived by correlating the first 77 quarters with the groups of 77 quarters which occur 1, 2, 3, .., 100 quarters later; in this way each correlation coefficient is based on 77 pairs of terms. The resulting curve B in fig. 6 contains the main features of A and

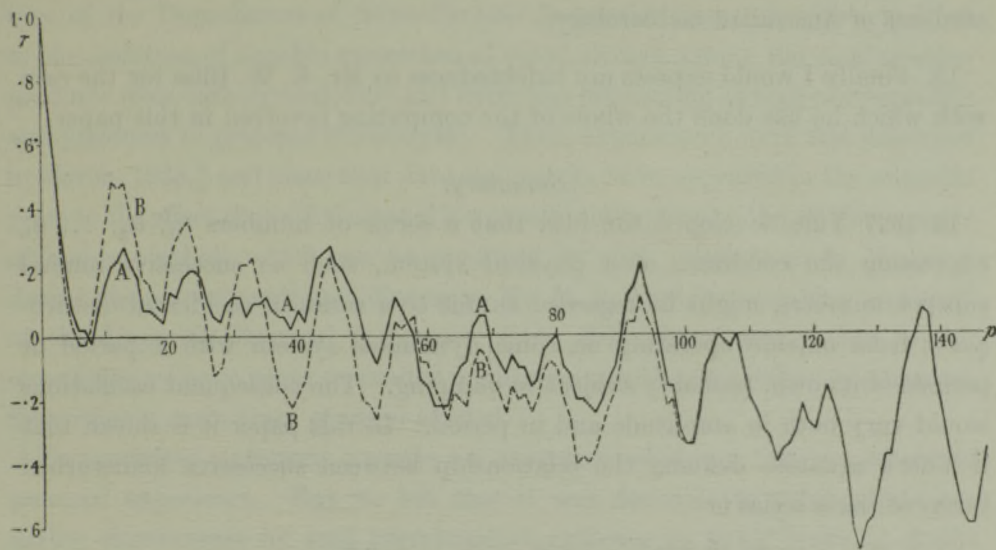

FIG. 6.—Serial Correlation Coefficients.

——, A, coefficients r_p based on all the available material. - - - - B, coefficients r_p based on 77 pairs of correlates.

shows the reduction in amplitude due to damping that we should expect. It appears wiser therefore to ignore values of r_p for which p is greater than about 100 in fig. 6; but the differences between curves A and B show that the error due to sampling (*i.e.*, to an inadequate number of pairs of terms) is probably as great as $0 \cdot 1$ and the apparent damping in B may be largely due to a greater amplitude in the $11\frac{1}{2}$ quarters oscillation in the first half of the data of 177 quarters than in the second.

14. On the whole then it is safer not to attempt a very precise interpretation. There appears to be a periodicity of about $11\frac{1}{2}$ quarters with an amplitude (or half-range) of something like $0 \cdot 12$, corresponding to an amplitude-ratio in the Fourier series of about $0 \cdot 35$; but the evidence that it is damped is not conclusive. There are also indications of a periodicity of about 46 quarters,

or $11\frac{1}{2}$ years, with an amplitude of the order of 0·05 (and amplitude ratio 0·22); of damping in this no trace is visible; this oscillation appears, therefore, to be superposed from without and is presumably solar in origin. The general downward slope of the graphs is due to the general trend of the pressure data. There is clear evidence of strong persistence.

The negative conclusions are more definite. Of the periodicities of 6·5 quarters and 13·5 quarters doubtfully suggested by the periodogram there is no visible trace; nor do I see any evidence of the other periods, such as 2 years, 4 years and 7 years, which have from time to time been suggested by students of Australian meteorology.

15. Finally I would express my indebtedness to Mr. E. W. Bliss for the care with which he has done the whole of the computing involved in this paper.

Summary.

In 1927 Yule developed the idea that a series of numbers $u_1, u_2, .., u_n$ expressing the condition of a physical system, such as successive annual sunspot numbers, might be regarded as due to a series of accidental disturbances from outside operating on some dynamical system with a period or periods of its own, probably subject to damping. The consequent oscillations would vary both in amplitude and in period. In this paper it is shown that if Yule's equation defining the relationship between successive undisturbed terms of the u series is

$$u_u = g_1 u_{x-1} + g_2 u_{x-2} + \dots + g_s u_{x-s},$$

then, provided n is large, a similar equation holds very approximately between successive values of r_p, the correlation coefficient between terms of u separated by p intervals, *i.e.*,

$$r_x = g_1 r_{x-1} + g_2 r_{x-2} + \dots + g_s r_{x-s}.$$

Thus the graph expressing the r_p's, which is much smoother than that of the u's, may be used to read off the character of the natural periods of the u's; further various relationships are found between the amplitude of the corresponding terms in the Fourier periods and those of the correlation coefficients.

The analysis is illustrated by applying it to the quarterly values of pressure at Port Darwin, a key-centre of world weather, which proves to have a strong persistence and to show evidence of not very strongly developed periods of about $34\frac{1}{2}$ months and of about four times this length or $11\frac{1}{2}$ years; the series of data is not long enough to settle whether the former oscillations are damped and are free oscillations, but the latter appear to be imposed from without and are presumably solar in origin.