De Restricted
~~CONFIDENTIAL~~

M.R.P. 118
2nd July, 1943.
~~SECRET~~

AIR MINISTRY
Meteorological Research Committee.

A second note on Dr. Brooks' method of long-range forecasting by means of pressure wave analysis.

by Sir Gilbert T. Walker.

When considering the efficiency of the difference periodogram as an instrument of research it seems desirable to supplement previous work by a more thorough examination of an interval of some months: and as we were provided in Dr Brooks' report with results given by Alter's method and by a mechanical analyser as well as by the difference periodogram for the interval of 144 days from 26.11.37 to 18.4.38 this seemed a suitable material for consideration.

The first step was to carry out a periodogram by harmonic analysis for periods of 3,4,5,6,7,8,9,10,11,12,13,14,16,18,20,22, 24,27,30,33,36,42,48 and 144 days. The values of the amplitudes in mbs. are indicated by the continuous line in Fig.1., and it will be noticed that they increase fairly consistently with the length of the period from 1 or 2 mb. for the shortest periods to 7 or 8 mb . for the longest. This feature has been noticed previously and I believe it is due to the close relationship between the pressures of successive days, the correlation coefficient r between them being ·8 or more. This is equivalent to a kind of stickiness or viscosity; and if we had two equal external controls, one with a period of 4 days and one of 40 days, the amplitude of the pressure oscillation set up in the former with its quick reversals would be much less than in the latter case, where the effects of the controls would be cumulative. The physical nature of the controls in pressure is very complex, and in order to get some idea of the effect of the stickiness I have, as before, considered the simple case in which u_s, the pressure of the s'th day, is related with that of the day before by

$$u_s = ru_{s-1} + v_s$$

where r = .83 and v_s is regarded as measuring the control. Let c be the amplitude of a period of u and f the ratio of c to $d\sqrt{2}$, where d is the standard deviation of u, so that for a pure sine curve f=1; and let f' be the corresponding amplitude-ratio of v. Then we have, as previously pointed out,

$$(f'/f)^2 = (1-2r\cos\beta + r^2)/(1-r^2)$$

where β is $360^\circ/p$, p being the number of days in the period. Typical ratios of f'/f for different values of p are:-

p	3	4	6	9	12	16	21	27	36	72	144	1000
f'/f	2.85	2.33	1.66	1.16	.90	.71	.57	.48	.41	33	30	30

The values of f' are shown by the dotted line in Fig.1 and it will be seen that the slope is very greatly diminished; in fact f' for p=144 is less than for p=3 and 4. The number n of terms included in the analysis varies from 126 when p=42 to 150 when p=15.

As the successive values of v are by hypothesis independent, we can apply to them Schuster's analysis of what accidental quantities will produce; thus for n days f' is as likely as not to exceed $1.18/n^{1/2}$ and when n=144 this is approximately .1. Such a value is called a 'probable' value and from Fig.1 it may be seen that only 14 out of 24 values of f' exceed .1. Accordingly this diagram provides little evidence of reality in the periods.

FG1A ~~S~~2

The probable value of $1.18/n^{\frac{1}{2}}$ for f' when the terms are fortuitous and the true value is zero has another interpretation. It is the error due to have n terms instead of an infinite series - in other words it is the 'error due to sampling'. If then we do not know that the terms are accidental, and our analysis gives, say f'=.5, the error due to sampling will be less than $1.18/n^{\frac{1}{2}}$ but will be of the same order of magnitude. It is easy enough to tabulate f' to two places of decimals, but the figure in the second place has not much value.

Further, a wave of p days may last for a time and re-appear with a nearly reversed phase, so that the contributions to the analysis of the whole interval of 144 days may largely cancel. It seemed desirable therefore to limit the number of days included in one analysis to 3p and to effect as many analyses as 144 days would provide, with extensions to 150 when necessary. The periods examined were twelve, of 5,6,7,8,9,10,11,12,13,18,22 and 27 days. It became clear that the mean pressures as well as the standard deviations d change materially from interval to interval, and those of the successive groups of 15 days are plotted in Fig.2, A and B: these changes have to be allowed for in each group. Thus for the period of 6 days (Fig.2, C) there are 8 groups of 18 days: for the third d= 15.8 mb, but for the fifth d=4.2, about a quarter of the former.* A constant amplitude of 5 mb. in the pressure wave would yield f'=.22 in the former group and would mean nothing, while in the later group f' would be about .8 and would be significant. Accordingly Dr.Brooks' method of basing the existence of periods on the number of days when the amplitudes are estimated to exceed 5 mb. has serious disadvantages

When showing in Fig.2, C the results of this analysis ≠ the values of a and b (the amplitudes of the cosine and sine terms in mbs.) have been plotted in order to show their amount of persistency. Those of c and f' are also given.

It is desirable to know whether any of the values of f' over an interval of 3p days are too big to be fortuitous. If we include the periods of 33,36,42 and 48 days in Fig.1 with the 64 in Fig.2 there are 72 values of f'; and if we pick out the biggest of 72 quantities of which each independently is as likely as not to exceed £, its value will be as likely as not to exceed 2.6£. Thus any value of f' is not significant unless it is materially larger than either .75 for p=5, .6 for p=9, .4 for p=18 or .33 for p=27. Of these limits perhaps .75 and .6 should be somewhat reduced owing to overestimation of the ratio f'/f when p is small: existing theory does not apply in such cases.

It will be noticed in Fig.2 that, as the theory of fortuitous variations would suggest, the values of f' diminish from p=5 to p=48. During the first 72 days, when the amplitudes of c are relatively large, none of the values of f' are materially in excess of the limits laid down except, perhaps, that of .8 for the 5 day period. It is in the second 72 days, when c is less than Dr.Brooks' 5 mb., that f'reaches .9 in our approximate calculation and is perhaps significant in the periods of 5 and 6 days. For 9 days f' is only a trifle larger than .6 and for 12 days it exceeds .4.

The conclusion reached by Dr. Brooks about the 144 days in question was based, I believe, on the magnitude of c. It was that there were periods of 6, 8½ to 9⅓, 17 or 18, and about 26 and 33 days; and that the period of 17 or 18 was specially well marked. Of these I regard my 27 as equivalent to his 26 because both lie in the interval between the subharmonics 144/6 and 144/5, and the resolving power is not sufficient to separate them.

* For the first half of the 144 days d= 15.4, and for the second half d=6.4.

≠ In the data for p=27 the first group contains 81 days, but for the second only 54 were available.

....

We may now turn to the curves of Alter and of the harmonic analysis. Alter's method is based on the correlation periodogram rather than on harmonic analysis, so that changes of phase are not brought to light; like Brooks' method it does not take the coefficient Υ into account. His curve has not very much in common with the results got by others. The graph from harmonic analysis resembles my curve for c, but there are some discrepancies; I cannot examine them as I have not been able to look up the original paper.

Happily in view of its difficulty, this question about the effects of accident can be seen from another angle. What really matters is whether a period of p days which is well marked over 3p days can be relied on to persist over another 3p days; and definite information about this is now available. Useful persistence depends on the signs of a and b; for a reversal means a forecast in the wrong direction. Now the graphs show 21 cases (see Table 1) in which over 3p days the amplitude of a or b is 4 mb. or over. During the following 3p days there are 5 cases (with periods 5,* 8* and 27 days) in which the amplitude was less than 2 mb, so that the wave had practically died out. Of those which had amplitudes from 2 to 3.9 mb. there were 5 in which the sign had persisted (of 6,7,9,10,18 days) and 4 in which it reversed (of 5,6,9 and 27 days): and of the 7 in which the succeeding amplitude was 4 mbs. or over, 3 had the same sign (9*and 12) and 4 the opposite (5,6* and 10). Hence as far as these 144 days are concerned persistence from one interval of 3p days to the next gives no better indication than tossing a penny.

It would not be fair to base a verdict regarding the validity of Dr.Brooks' methods on this examination of a single interval of time: but it leaves me with a strong impression that it is not worth while to tackle these difficult questions by any method that does not give quantitative results, or in which the truth of a principle cannot be estimated by definite numbers of its successes and failures on a scale large enough to provide reliability.

* occurring twice.

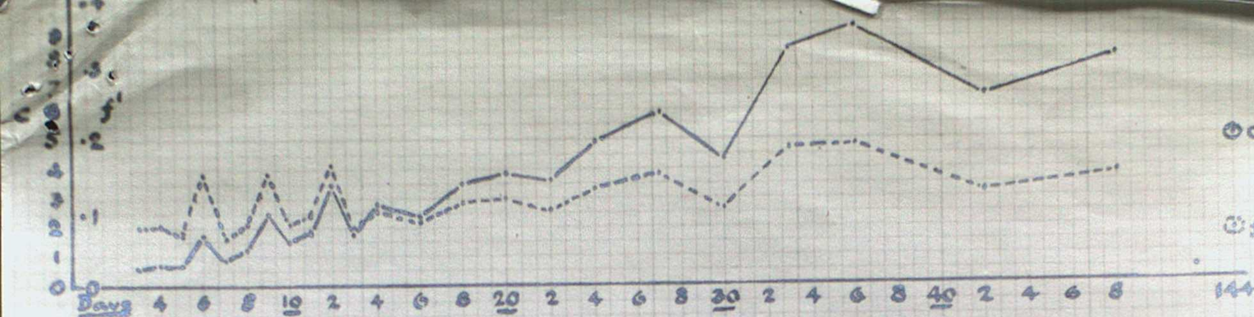

Fig 1. Values of c and f' for different periods.

Fig. 2