M.O. 420 q AIR MINISTRY

METEOROLOGICAL OFFICE

PROFESSIONAL NOTES NO. 97

(*Seventeenth Number of Volume VI*)

SOME CORRELATION COEFFICIENTS BETWEEN CERTAIN UPPER AIR DATA

BY SIR GILBERT T. WALKER, C.S.I., F.R.S.

LONDON
HIS MAJESTY'S STATIONERY OFFICE
1949

TWOPENCE NET

Decimal Index
551.506.7 : 519.2

Prof. Notes Met. Off.
London 6, No. 97, 1949

SOME CORRELATION COEFFICIENTS BETWEEN CERTAIN UPPER AIR DATA

By Sir Gilbert T. Walker, C.S.I., F.R.S.*

I give below an account of results got from radio-sonde data of June–October 1941, and have prefaced it by a summary of previous conclusions derived from data sent me by Dr. Dobson at various times.

The following symbols are used :—

O = Ozone value measured by Dobson's method.

P_n = Pressure at a height of n Km.

T_n = Temperature at a height of n Km.

H_c = Height of tropopause.

(xy) = Correlation coefficient between any variables x and y.

R = Multiple correlation coefficient of a regression equation.

Data of O, P_5, and T_7 from January to September 1940 led to the correlation coefficients in Table I in which values of the previous day are distinguished by dashes.

TABLE I—CORRELATION COEFFICIENTS MULTIPLIED BY 100

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Jan.–Sept.
Ozone and pressure at 5 Km.										
(OP_5)	−45	−68	−67	−65	−14	−76	−74	−03	−65	−54
$(O'P_5)$	−36	−53	−45	−51	−28	−49	−54	06	−52	−45
(OP_5')	−46	−51	−51	−55	−01	−67	−74	06	−55	−42
Ozone and temperature at 7 Km.										
(OT_7)	−34	−73	−66	−65	−64	−65	−88	03	−85	−59
$(O'T_7)$	−12	−64	−34	−45	−57	−41	−60	−14	−62	−41
(OT_7')	−43	−66	−26	−41	−26	−54	−77	−24	−36	−46
Pressure at 5 Km. and temperature at 7 Km.										
(P_5T_7)	71	80	64	81	52	75	79	80	87	73
$(P_5'T_7)$	55	49	54	74	32	55	70	70	62	66
(P_5T_7')	64	57	26	62	44	62	60	68	68	52
Persistence (between values on consecutive days)										
(OO')	75	79	52	48	76	68	79	70	67	68
(P_5P_5')	83	50	78	74	73	79	83	79	72	77
(T_7T_7')	85	61	39	73	57	77	68	73	70	70

My chief aim at this time was to get light on the physical priorities—whether, for instance, ozone controlled the temperature so far in advance that O was more closely related with P_5 or T_7 of the following day than it was with P_5 or T_7 of the same or previous days.

---

* The following paper consists essentially of matter prepared for a research committee and was written when upper air information was not as reliable as it is today. It is very probable therefore that some of the numerical results will require modification. G.T.W., 1948.

The last column in Table I with $-\cdot 54$ and $-\cdot 59$ compared with $-\cdot 45$ and $-\cdot 41$ shows that this is not the case. Also P_5 and T_7 are more nearly related to each other than either is to O.

The variability in the monthly values of the correlation coefficients in Table I is large ; a month is not nearly long enough to give reliability.

In May 1941 Dr. Dobson gave me data for 33 days for O and H_c, T_7 and P_0 at Boscombe Down in the hope that it might be possible to deduce the height of the tropopause from the other data. Conclusions based on so short a period, with only 27 pairs of consecutive days, clearly need confirmation.

The correlation coefficients, multiplied by 100, were :—

	H_c	O	T_7
O	−68	—	
T_7	76	−48	—
P_0	60	−50	36

Here we have $(H_cO) = -\cdot 68$. On the other hand, between H_c as observed at Boscombe Down and at Penzance the correlation coefficient is $\cdot 55$; so that O gives a better indication for H_c at Boscombe Down than does H_c observed at Penzance. For T_7 we find $(T_7H_c) \doteq \cdot 76$. If in future h, w, t, p, stand for the departures H_c, O, T and P divided by the standard deviations, we also find

$$h = -\cdot 41\,w + \cdot 56\,t_7 \qquad \text{with } R = \cdot 84, \qquad \ldots\ldots \quad (1)$$

$$h = \cdot 63\,t_7 + \cdot 37\,p_0 \qquad \text{with } R = \cdot 76, \qquad \ldots\ldots \quad (2)$$

$$h = -\cdot 30\,w + \cdot 52\,t_7 + \cdot 26\,p_0 \qquad \text{with } R = \cdot 87. \qquad \ldots\ldots \quad (3)$$

I was later given the data for P_7 during the same period, and the addition led to a material improvement. The correlation coefficients of P_7 with H_c, O and T_7 are $\cdot 82$, $-\cdot 51$ and $\cdot 72$ as compared with $\cdot 60$, $-\cdot 50$ and $\cdot 36$ for P_0. Instead of (3) we get

$$h = -\cdot 30w + \cdot 29\,t_7 + \cdot 46p_7 \qquad \text{with } R = \cdot 89. \qquad \ldots\ldots \quad (4)$$

If data of the previous day, distinguished by dashes, alone were available we should have

$$h = \cdot 05w' + \cdot 76p_7' - \cdot 02t_7' \qquad \text{with } R = \cdot 71. \qquad \ldots\ldots \quad (5)$$

As pressure of the previous day seemed of outstanding importance I formed the equation

$$h = -\cdot 30w + \cdot 12p_7 + \cdot 41t_7 + \cdot 36p_7' \text{ with } R = \cdot 92, \qquad \ldots\ldots \quad (6)$$

a better result than that from (4).

A regression equation in terms of p_7' and t_7 yields $R = \cdot 85$, while one in terms of p_7 and t_7 yields $R = \cdot 75$. One for h in terms of w and p_7' leads to $R = \cdot 82$. The value w' of ozone on the previous day is found to be insignificant.

In November 1941 two questions on the determination of H_c from values of pressure and temperature were raised :—

(*a*) whether measurements at 5 Km. were comparable in value with those at 7 Km. ? and

(*b*) whether any use could be made of P_0 and T_1 ?

Regarding (*a*) it appears from 5 months of radio-sonde information that data from 5 Km. are very nearly as useful as those from 7 Km. The correlation coefficients are :—

	P_7	T_7
H_c	·81	·78
P_7	—	·84

	P_5	T_5
H_c	·79	·74
P_5	—	·85

and, while the coefficient between H_c as reported and as calculated is ·82 for 7 Km., it is ·80 for 5 Km. Tabulation over June and July shows an average error of 1,720 ft. in H_c as determined from P_5 and T_5, the standard deviation from June to October being 5,085 ft.

Regarding (*b*) I reported in January 1942 that the regression equation of H_c in terms of P_0 and T_1 led to a result for which $R = \cdot 74$.

Later in January, I described some vain efforts to trace any method in the variations of the daily values of H_c regarded as a series of terms ; I could find no appreciable relationship beyond mere persistence, (hh') being ·42. The irregularity in the changes of H_c was noted by Dines in 1911* ; he remarks : " There is evidence that, in the 40 miles between Pyrton Hill and Ditcham Park, it may differ by at least 1 Km., and that the same change may occur in the same locality in less than an hour."

Further work has led to the following table of correlation coefficients between contemporary values :—

TABLE II—CORRELATION COEFFICIENTS MULTIPLIED BY 100

	H_c	O	P_7	P_0	T_7
O	−46	—			
P_7	81	−44	—		
P_0	64	−41	69	—	
T_7	78	−38	89	38	—
T_1	46	−41	71	12	69

With factors measured at 5 Km. the relations are, as stated before, very nearly equal to those at 7 Km.

The correlation coefficients are in general agreement with those found by others, the chief difference being that here O appears to be less closely related

* DINES, W. H. ; The vertical temperature distribution in the atmosphere over England and some remarks on the general and local circulations. *Philos. Trans., London*, A, **211**, 1911, p. 253.

with the pressure and temperature. While the bigness of such a coefficient as ·89 between P_7 and T_7 is partly due to the mere process of working out P and T from the observations, other fairly high values are independent of such considerations and inspire confidence in the reliability of the data.

For relationships between values of factors on any day and values on the previous day (the latter being dashed) we have :—

$$\begin{array}{lll} (hp_7') = \cdot 42, & (h'p_7) = \cdot 62, & \text{while } (hp_7) = \cdot 81 \\ (hp_0') = \cdot 43, & (h'p_0) = \cdot 33, & \text{while } (hp_0) = \cdot 64 \\ (ht_7') = \cdot 32, & (h't_7) = \cdot 29, & \text{while } (ht_7) = \cdot 78 \\ (p_7p_0') = \cdot 37, & (p_7'p_0) = \cdot 30, & \text{while } (p_7p_0) = \cdot 69 \\ (p_7t_7') = \cdot 42, & (p_7't_7) = \cdot 47, & \text{while } (p_7t_7) = \cdot 89 \\ (p_0t_7') = \cdot 13, & (p_0't_7) = \cdot 24, & \text{while } (p_0t_7) = \cdot 38 \\ (p_0t_1') = -\cdot 06, & (p_0't_1) = \cdot 14, & \text{while } (p_0t_1) = \cdot 12 \end{array}$$

It will be noticed that correlations with p_0 on the previous day are somewhat larger than those with p_0 on the following day. This does not apply to p_7.

Denoting the change in a day by Δ, so that $\Delta h = h_m - h_{m-1}$, and $\Delta h' = h_{m-1} - h_{m-2}$, we find

$$(\Delta h \Delta p_0') = -\cdot 17, \quad (\Delta h' \Delta p_0) = -\cdot 28, \quad \text{while } (\Delta h \Delta p_0) = \cdot 52.$$

Also $(\Delta h \Delta h') = -\cdot 31.$

If changes were due to advection of large masses we should expect appreciable continuity in the changes from one day to the next ; so that the value of $-\cdot 31$ indicates considerable patchiness or variability in the horizontal distributions (see p. 4). If the correlation coefficient between the values of h separated by n days is denoted by r_n we find

$$r_1 = \cdot 42, \quad r_2 = \cdot 16, \quad r_3 = \cdot 14, \quad r_4 = \cdot 23, \quad r_5 = \cdot 27.$$

The regression equations worked out are :—

$$h = \cdot 56p_7 + \cdot 28t_7, \quad \text{with } R = \cdot 82 \quad \ldots\ldots (7)$$

$$h = \cdot 57p_5 + \cdot 26t_5, \quad \text{with } R = \cdot 80 \quad \ldots\ldots (8)$$

$$h = \cdot 64p_7' - \cdot 25t_7', \quad \text{with } R = \cdot 44 \quad \ldots\ldots (9)$$

$$h = \cdot 30h' + \cdot 47p_7' - \cdot 34t_7', \quad \text{with } R = \cdot 46 \quad \ldots\ldots (10)$$

The more comprehensive equation is

$$h = 1\cdot 1p_7 - \cdot 2t_7 - \cdot 3h' - \cdot 1p_7' + \cdot 2t_7', \text{with } R = \cdot 85 . \quad \ldots\ldots (11)$$

But such large changes are produced in the coefficients of equation (11) by small errors in the total correlation coefficients that conclusions cannot be very safely based on it.

When we attempt a physical interpretation of the relations in Table II we naturally consider the changes occurring during the passage of cyclones and anticyclones ; and of these a recent summary will be found in a paper by B. and E. Haurwitz.* Now the features of pressure at all heights associated with the passage of isallobaric highs and lows (*loc. cit.* Fig. 19, p. 49) agree fairly well with those in Table II. The features of temperature at 6–8 Km. agree also fairly well ; but those near the ground differ fundamentally, the temperature and pressure varying clearly in opposite directions, obviously having a large negative correlation coefficient, while in Table II $(p_0\, t_1) = \cdot 12$ in agreement with Dines' value $(p_0\, t_0) = \cdot 11$.

---

* HAURWITZ, B., and HAURWITZ, E. : Pressure and temperature variations in the free atmosphere over Boston. *Harvard met. Studies, Cambridge, Mass.*, No. 3, 1939.

Some such discrepancy is to be expected since in England the passage of typical recently formed cyclones is comparatively rare, our weather being mainly controlled by fronts of various types. Also the relations holding during the actual passage of the troughs and crests are likely to be swamped by the relations of the intervening periods.

Now if for any period of time we knew that the Haurwitz diagram would be valid in England, its chief use from the present standpoint would be that, owing to the westward inclination of the axis of a cyclone, changes in p_0 would precede changes in p_7 and t_7 and would give indications in advance of changes in the height of the tropopause.

It is natural, therefore, to see what has happened on the occurrence of a marked trough or crest of pressure, as shown on a graph of the conditions during the five months under examination. Of these events there have been about 35 ; and while in 18 of these the change in p_0 occurred on the same day as the changes at 7 Km., there were 15 in which p_0 was one day in advance, one in which p_0 was a day late, and one which was doubtful. These numbers are only approximate but their implications seem clear.

On the whole it must be admitted that the coefficients of Table II and p. 5 do not encourage the idea that in ordinary weather H_e can be successfully forecast a day ahead. But there is ground for hoping that when a crest or trough of pressure is travelling across the country the marked change in H_e associated with it may give useful information. Further it may be that the priority of p_0 is of significance.

I propose now to make further use of the data by separating the times at which troughs or crests were passing over England and finding statistically what are the relationships (*a*) at these times, and (*b*) during the intervening periods. Also, the data should, if possible, be corrected for seasonal changes ; for O_3 this has been done, but data of H_e, P, T have not been corrected. Some of the normals published by Dines have been violently smoothed and do not agree with those in the " Handbuch der Klimatologie "*

ADDENDUM 1942

In continuation, I have attempted to distinguish between relationships holding during weather of different types ; and as the factors studied have been H_e (the height of the tropopause) and P_0 (the pressure at ground level) for which fairly satisfactory monthly means are available I have corrected for seasonal changes.

A further examination of the graphs brought out a well marked feature that I had previously failed to notice. When the pressure changes are not too rapid the changes in P_0 tend to precede those of H_e ; but when pressure drops rapidly in 24 hours and then recovers rapidly, so that a narrow trough of low pressure passes in two days, there is in general no suggestion of precedence in P_0. Thus, June 1941 was a period of fairly large pressure variations without any very steep changes ; and while the correlation coefficient between P_0 and H_e for the same day, (P_0H_e), was ·61, that between P_0 and H_e of the day after, $(P_0'H_e)$, was ·70 and that between P_0 and H_e of the previous day, (P_0H_e'), ·42 only.

During the period of 5 months I picked out 7 periods of exceptionally narrow and steep fronts, as shown in the graph of P_0 : July 16–21, August 4–6,

---

* KÖPPEN, W., and GEIGER, R. : Handbuch der Klimatologie, Band I, Teil F, Berlin, 1931.

August 10–12, August 17–9, August 29–31, September 10–12, October 14–7, making 25 days in all. In most of these the group consists of three days, the day of lowest pressure and the days before and after. These periods yield

$$(P_0H_c) = \cdot 7, \qquad (P_0'H_c) = -\cdot 1, \qquad (P_0H_c') = \cdot 2.$$

On the other hand for the 116 days without narrow fronts

$$(P_0H_c) = \cdot 47, \qquad (P_0'H_c) = \cdot 52, \qquad (P_0H_c') = \cdot 21\,;$$

and for the whole period of 141 days under examination the coefficients are

$$(P_0H_c) = \cdot 53, \qquad (P_0'H_c) = \cdot 35, \qquad (P_0H_c') = \cdot 20.$$

Thus, while in the 116 days the correlation coefficient between P_0 and H_c of the following day exceeds the correlation coefficient for the same day, on days with narrow fronts there is no warning of the drop in H_c to be derived from the changes in P_0.

It appears therefore that on about 80 per cent. of days a better indication of H_c can be derived from P_0 of the day before than from that of the day itself. But the relationship of ·52 is not, by itself, big enough to have practical value ; it is necessary to use a formula involving data of the day in question.

The effect of correcting for seasonal changes is appreciable. For H_c I used Dines' data and for P_0 those of Greenwich. Uncorrected (P_0H_c) was ·64 ; corrected it is ·53.

(86749) Wt. 3853/M.1091 4/49 Hw. PRINTED IN GREAT BRITAIN

LONDON
PRINTED AND PUBLISHED BY HIS MAJESTY'S STATIONERY OFFICE
To be purchased directly from H.M. Stationery Office at the following addresses:
York House, Kingsway, London, W.C.2; 13a Castle Street, Edinburgh, 2;
39 King Street, Manchester, 2; 2 Edmund Street, Birmingham, 3;
1 St. Andrew's Crescent, Cardiff; Tower Lane, Bristol, 1;
80 Chichester Street, Belfast
OR THROUGH ANY BOOKSELLER
1949
Price 2*d.* net

40-9999