"On the Influence of the Time Factor on the Correlation between the Barometric Heights at Stations more than 1000 Miles apart." By F. E. CAVE-BROWNE-CAVE, Girton College, Cambridge. Communicated by Professor KARL PEARSON, F.R.S. Received June 3, and in revised form, December 20, 1904,—Read June 16, 1904.

1. An investigation of the relationship of the daily barometric heights on both sides of the Atlantic Ocean has been in progress for some years, and in a preliminary note by Professor Pearson and myself* some account was given of the contemporaneous relationship of a chain of stations from the extreme north of Norway down the west coasts of Europe and Africa. Observations for this east side of the Atlantic have now been copied for twenty years, as far as stations are available from Norway to the Cape, and the only need here is more aid in the very laborious reductions necessary before any inferences can be drawn. A similar chain of stations from Nova Scotia to the Falkland Isles has been completed, with the exception of Brazil, from which, so far, we have been able to obtain no data whatever.

2. It seemed desirable, before entering on the cross Atlantic correlations, to obtain a better conception of how the time factor influences the intensity of correlation. For this purpose two stations, Wilmington and Halifax, seemed specially appropriate. They are both on the East Atlantic coast, the one in North Carolina and the other in Nova Scotia. The results obtained for these stations afford an illustration of the manner in which the theory of correlation can be used for the purposes of meteorological prediction. In a paper published in 1898,† Professor Pearson and Dr. Lee indicated that the proper way to deal with the pressure relationship between two or more stations was to proceed by the method of correlation; and they illustrated this by the simultaneous correlations of various stations in the British Isles. The corresponding simultaneous correlations for stations from Norway to Sierra Leone were given by Professor Pearson and myself in the note already referred to. The object of the present paper is to illustrate this method of correlation as applied to prediction, the fundamental points being (*a*) the choice of two stations which, although far apart, have a considerable correlation with at least 12, and if possible, 24 hours' interval; (*b*) the determination for any two stations of the interval for which the correlation is a maximum. If the practical meteorologist is accustomed to predict roughly the barometric condition at a station A, from stations B, C, and D, then, if he has observations for twenty years or more for these stations, modern statistical methods place him in a position to choose

* 'Roy. Soc. Proc.,' vol. 70, pp. 465—470.

† 'Phil. Trans.,' A, vol. 190, pp. 423—469.

the best intervals between the readings in order to find, from the anterior observations at the four stations, the most probable barometric height at A on a given day, and the most probable error of this estimate. They thus supply a means of giving to his predictions a greater precision than is at present attainable. It should also be remembered that, although the present paper only deals with barometric observations, the same methods could be applied to temperature, or to any other of the quantities with which meteorology is concerned.

In order to show that the relations between the two stations mainly considered in this paper are not due to purely local conditions, corresponding investigations have been made for two stations in the Southern Hemisphere, St. Helena and Cape Town. Although the shortness of the period (1893–98) for which observations taken at St. Helena are available, prevents these results from being as fully reliable as those obtained for the North American stations, yet they are sufficient to show the general nature of the pressure relationship between the two places, and the influence of the time factor.

The latitudes and longitudes, and hours of observation are as follows :—*

Halifax, 44° 35′ N., 63° 40′ W.; 9 A.M.

Wilmington, 33° 14′ N., 77° 59′ W.; 7 A.M. up to June 30, 1888, afterwards 8 A.M.

Cape Town, 34° 0′ S., 18° 35′ E.; 8 A.M.

St. Helena, 15° 57′ S., 5° 40′ W.; 9 A.M.

This gives an arc of about 15° 51′, or about 1090 miles between Halifax and Wilmington, and 28° 18′, or about 1950 miles between Cape Town and St. Helena. Roughly speaking, Halifax is north-east of Wilmington, and Cape Town south-east of St. Helena.

The twenty years of the data for Halifax and Wilmington were divided into two decades, 1879–88 and 1889–98, and the summer and winter months (equinox to equinox) were dealt with separately. As was pointed out in the previous note, there is a marked difference between the summer and winter results, and two ten-year periods were desirable for mutual control. It soon became obvious that the summer and winter time factors were different in character, the interval for maximum correlation being considerably shorter in the summer. The numbers obtained indicate that change of pressure passes from Wilmington to Halifax, *i.e.*, from south-west to north-east, and thus it

* The Canadian Meteorological Office kindly superintended the copying of the Halifax and other Canadian data, the cost of which was defrayed by Professor Pearson; the Smithsonian Institution provided the copies of the Wilmington and other United States data; the readings for Cape Town we owe to the mediation of Dr. W. N. Shaw, who procured them from the Cape Meteorological Committee, the cost of copying them being defrayed by a grant from the Government Grant Committee of the Royal Society.

is desirable to take Wilmington earlier in order to predict for Halifax later. For the southern pair of stations there is the same kind of difference between the summer and winter results; this seems to indicate that the phenomenon is really connected with the season, and does not depend on influences external to the Earth, since the summer for the first pair is the winter for the second. Moreover, it is found that for the two southern stations the change of pressure passes from north-west to south-east, so that in predicting we ought to take St. Helena earlier than Cape Town. The following tables give the correlations obtained:—

Table I.—Correlation of Barometric Heights at Halifax and Wilmington.

Interval between readings.*	Summer, 1879—1888.	Winter, 1879—1888.	Interval between readings.	Summer, 1889—1898.	Winter, 1889—1898.
hrs.			hrs.		
−118	+0·0585	—	−119	−0·0134	—
− 94	+0·0605	—	− 95	−0·0247	—
− 70	+0·0668	+0·0496	− 71	−0·0141	+0·0512
− 46	+0·1391	+0·0379	− 47	+0·0418	+0·0598
− 22	+0·2385	+0·0416	− 23	+0·1742	+0·0420
+ 2	+0·3432	+0·2878	+ 1	+0·3038	+0·2456
+ 26	+0·3714	+0·5206	+ 25	+0·3143	+0·4701
+ 50	+0·2176	+0·2344	+ 49	+0·1538	+0·2263
+ 74	+0·1138	+0·0842	+ 73	+0·0581	+0·0839
+ 98	+0·0816	+0·0342	+ 97	+0·0150	+0·0736
+122	+0·0881	+0·0217	+121	−0·0050	+0·0881

Table II.—Correlation of Barometric Heights at Cape Town and St. Helena.

Interval between readings.†	Summer, 1893—1898.	Winter, 1893—1898.
hrs.		
−97	+0·2202	+0·2898
−73	+0·2278	+0·2470
−49	+0·2620	+0·2405
−25	+0·3461	+0·2919
− 1	+0·4478	+0·3927
+23	+0·4468	+0·4605
+47	+0·3434	+0·4081
+71	+0·2908	+0·3409
+95	+0·3033	+0·3130
+119	+0·2950	+0·2785

* Positive when the reading is taken later at Halifax.

† Positive when the reading is taken later at Cape Town.

3. It seemed desirable to determine the intervals for which the correlation between each pair of stations is greatest, and also the values of these maxima. After some consideration, the curve

$$\log y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

was selected as likely to give a reasonable interpolation result,* the axis of y being that of the correlation, and x being the interval, in days, between the readings, counted as positive when the observation is taken later at the more easterly station in each case. The curves are not intended to represent the complete correlation data, but they have been used as convenient empirical methods to find in position and magnitude the maximum ordinates. For this reason only five correlations were used for calculating each curve, namely, those given in rows 5 to 9 of Table I, and rows 4 to 8 of Table II. The following curves were determined :—

WILMINGTON AND HALIFAX.

Summer, 1879—1888 :—

$$\log_{10} y = -0{\cdot}477451 + 0{\cdot}161608x - 0{\cdot}062808x^2 - 0{\cdot}058772x^3 + 0{\cdot}014994x^4.$$

Maximum ordinate: 0·3877 at x = 0·7275, *i.e.*, 17·46 hours.

Summer, 1889—1898 :—

$$\log_{10} y = -0{\cdot}524516 + 0{\cdot}175661x - 0{\cdot}122239x^2 - 0{\cdot}040868x^3 + 0{\cdot}012096x^4.$$

Maximum ordinate : 0·3386 at x = 0·5862, *i.e.*, 14·07 hours.

Winter, 1879—1888 :—

$$\log_{10} y = -0{\cdot}592740 + 0{\cdot}647374x - 0{\cdot}300470x^2 - {\cdot}054822x^3 + 0{\cdot}021968x^4.$$

Maximum ordinate: 0·5272 at x = 0·9553, *i.e.*, 22·92 hours.

Winter, 1889—1898 :—

$$\log_{10} y = -0{\cdot}634964 + 0{\cdot}615514x - 0{\cdot}258629x^2 - 0{\cdot}073287x^3 + 0{\cdot}025021x^4.$$

Maximum ordinate: 0·4721 at x = 0·9674, *i.e.*, 23·22 hours.

ST. HELENA AND CAPE TOWN.

Summer, 1893—1898 :—

$$\log_{10} y = -0{\cdot}346171 + 0{\cdot}063222x - 0{\cdot}064491x^2 - 0{\cdot}011993x^3 + 0{\cdot}006497x^4.$$

Maximum ordinate: 0·4660 at x = 0·4518, *i.e.*, 10·84 hours.

Winter, 1893—1898 :—

$$\log_{10} y = -0{\cdot}400900 + 0{\cdot}119333x - 0{\cdot}039277x^2 - 0{\cdot}022393x^3 + 0{\cdot}006580x^4.$$

Maximum ordinate : 0·4596 at x = 0·9994, *i.e.*, 23·986 hours.

* The suggestion is due to Mr. W. Palin Elderton.

4. It therefore appears that the maximum correlation between Halifax and Wilmington in the summer occurs when the barometer at Halifax is read about 16 hours later than that at Wilmington, and that the corresponding interval for the winter is about 23 hours. The intervals for St. Helena and Cape Town are about 11 hours and 24 hours respectively. Some caution may be necessary in accepting the values obtained for the summer intervals, since it is possible that the diurnal variation may affect the correlation; this point could only be settled by means of observations taken at other hours of the day. This effect of the diurnal variation, if it really exists, may conceivably account for some part of the difference between the intervals obtained for the summer months of the two decades for the American stations. Nevertheless, it seems reasonable to conclude that, in order to predict the barometric height at Halifax at 9 A.M., it is desirable to use the reading taken at Wilmington at 10 A.M. on the previous morning in the winter, while in the summer the Wilmington reading should be taken at about 5 P.M.

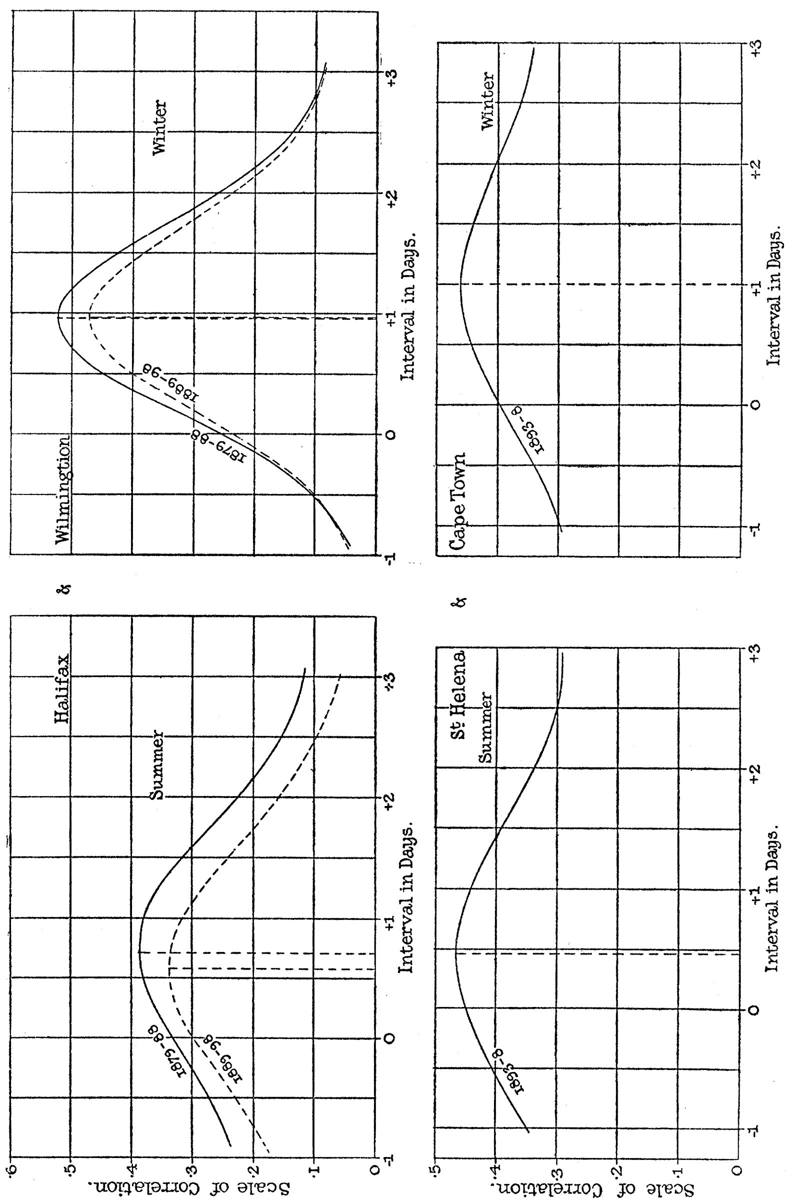
Diagram I (next page) shows the interpolation curves used.

5. The physical interpretation of these intervals for which the correlation is a maximum, is not obvious to me, although it may possibly be so to practical meteorologists. Of course, the intervals, 11 to 24 hours, are too short to be looked upon as corresponding to the transfer of an actual disturbance between the stations. Even if this were not the case, they could not give the average intervals between the arrival of the same isobar at the two places considered, since none of the pressures correlated may be alike. I should be glad to see the point adequately treated by one with practical meteorological experience, and will merely remark here that the correlation depends not upon equality of pressure, but upon proportionality of deviations from the local means. A full explanation must also account for the fact that the winter and summer relations differ in a very marked way.

6. Another point which seemed to deserve investigation was the correlation between the daily rise or fall of the barometer at Halifax and Wilmington. This investigation would have involved considerable labour if it had been necessary to proceed by the direct method of preparing tables of the daily rise or fall, and then calculating the correlations from them. The work has been considerably lessened by the use of the following method, which is due to Professor Pearson.

Let x_1 be the reading on one day and x_2 that on the following day at the first station, x_1' and x_2' the corresponding readings at the second station after any interval. Let m_1, m_2, m_1', m_2' be the means, σ_1, σ_2, σ_1', σ_2' the standard deviations, r_{12} the correlation of x_1 and x_2, $r_{1'2'}$, that of x_1' and x_2', $r_{12'}$, that of x_1 and $x_{2'}$, and so on, and R that of $(x_1 - x_2)$ and $(x_1' - x_2')$. The mean values of $(x_1 - x_2)$ and $(x_1 - x_2')$ will

DIAGRAM I.



clearly be $m_1 - m_2$ and $m_1' - m_2'$ respectively. Hence, if N be the number of pairs correlated, and if S denotes summation, we have

$$R = \frac{S\{(x_1 - x_2) - (m_1 - m_2)\}\{(x_1' - x_2') - (m_1' - m_2')\}}{N\sqrt{S\{(x_1 - x_2) - (m_1 - m_2)\}^2}\sqrt{S\{(x_1' - x_2') - (m_1' - m_2')\}^2}}$$

$$= \frac{\sigma_1\sigma_1' r_{11'} - \sigma_2\sigma_1' r_{21'} - \sigma_1\sigma_2' r_{12'} + \sigma_2\sigma_2' r_{22'}}{\sqrt{(\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2 r_{12})(\sigma_1'^2 + \sigma_2'^2 - 2\sigma_1'\sigma_2' r_{1'2'})}} \quad \ldots\ldots\ldots\ldots$$

It may be noted that this formula would apply to any case in which it was desired to correlate the difference of one pair of quantities with the difference of another pair. If we were not dealing with the summer and winter months separately, we should have, with sufficient approximation,

$$\sigma_1 = \sigma_2, \qquad \sigma_1' = \sigma_2', \qquad r_{11'} = r_{22},$$

since the quantities included in the x_1 and the x_2 groups would be the same, with the exception of a single reading at each end of the period considered. The formula would then take the simple form

$$R = \frac{2r_{11'} - r_{21'} - r_{12'}}{2\sqrt{(1 - r_{12})(1 - r_{1'2'})}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}.$$

But in our case the x_1 and x_2 groups differ by a reading at each equinox of each year considered, and it was therefore decided to use the accurate equation (i), although (ii) would have been sufficient to give a good idea of the correlation.

In order to use this formula, it was necessary to calculate the correlations between the barometric heights on consecutive days at Halifax and at Wilmington. Table III gives these correlations, and Table IV gives the correlations between the daily rise at the two stations for various intervals.

Table III.—Correlation of Barometric Heights on Consecutive Days.

Station.	Summer, 1879—1888.	Winter, 1879—1888.	Summer, 1889—1898.	Winter, 1889—1898.
Halifax	0·5924	0·3998	0·5155	0·4270
Wilmington	0·6151	0·4827	0·6510	0·4916

The most noticeable fact which appears from these results is the analogy between them and the space relations considered in the previous note. In that case we found that with increasing distance the correlation diminished to a negative minimum, and then began to increase; and the negative correlations obtained were of considerable magnitude. This does not seem to occur for the time relations when the actual readings are considered; for, although the values given in Tables I and II seem to indicate that a minimum has been reached, the only negative correlations which have been found are so small as to be

Table IV.—Correlation of Daily Rise at Halifax and Wilmington.

Interval between readings.*	Summer, 1879—1888.	Winter, 1879—1888.	Interval between readings.	Summer, 1889—1898.	Winter, 1889—1898.
hrs.			hrs.		
−94	−0·0071	—	−95	−0·0349	—
−70	−0·0801	—	−71	−0·0540	—
−46	−0·0298	−0·0166	−47	−0·0899	+0·0241
−22	−0·0107	−0·2152	−23	+0·0034	−0·2047
+ 2	+0·0960	+0·0133	+ 1	+0·1450	−0·0200
+26	+0·2268	+0·4676	+25	+0·2043	+0·4353
+50	−0·0669	−0·1205	+49	−0·0775	−0·0950
+74	−0·0919	−0·0901	+73	−0·0638	−0·1227
+98	−0·0391	−0·0367	+97	−0·0199	−0·0260

of very little significance. But it does occur in a very marked way for the time relations when we correlate the daily rise at the two stations. This analogy strengthens our belief that the intervals of maximum correlation are of great importance from the standpoint of atmospherical physics, and our desire to see the practical meteorologist's explanation of the matter is correspondingly increased.

7. The magnitude of the correlations which have been found to exist between the readings on consecutive days at Wilmington and Halifax, suggests the possibility of predicting barometric heights at the latter station. with a fair degree of accuracy from the heights observed at Wilmington on the previous day. In the paper by Professor Pearson and Dr. Lee already referred to, a table is given showing the barometric heights at Stonyhurst on 50 days, calculated from the simultaneous readings at Southampton and Laudale. The difference between the calculated and observed heights at Stonyhurst was small in every case, the mean value being about $\frac{1}{40}''$. At that time no correlations had been calculated between observations taken on different days; but the values now obtained make it possible to apply the theory to actual prediction. In predicting for Halifax from a single station more than 1000 miles away, we are naturally unable to obtain such good results, but, on the other hand, we have the advantage of calculating the probable heights at Halifax a day beforehand. In order to see how far the prediction is verified, forty dates were taken at random, one in summer, and one in winter for each year. It was impossible to allow the best intervals, 16 hours in summer, and 23 in winter, as we do not possess readings for Wilmington taken at the required times, and the predictions have, therefore, been made with the interval of 26 hours for the first 10 years, and 25 for the second.

* Positive when the readings are taken later at Halifax.

Table V gives the means, standard deviations, and regression coefficients which were employed in the calculations.

Table V.—Barometric Constants for Halifax and Wilmington.

Station and season.	Mean.	Standard deviation.	Regression.
Halifax, summer, 1879—1888	29″·842	0″·2446	0·6096
Wilmington, do.	30·045	0·1490	
Halifax, summer, 1889—1898	29·870	0·2215	0·4681
Wilmington, do.	30·085	0·1487	
Halifax, winter, 1879—1888	29·882	0·3446	0·8764
Wilmington, do.	30·152	0·2047	
Halifax, winter, 1889—1898	29·868	0·3436	0·7678
Wilmington, do.	30·158	0·2104	

The corresponding prediction equations are as follows, if H_p denote the probable height at Halifax corresponding to a height W observed at Wilmington on the previous day.

Summer 1879—1888 .... $H_p = 11{\cdot}527 + 0{\cdot}6096W$.
Summer 1889—1898 .... $H_p = 15{\cdot}787 + 0{\cdot}4681W$.
Winter 1879—1888........... $H_p = 3{\cdot}457 + 0{\cdot}8764W$.
Winter 1889—1898........... $H_p = 6{\cdot}713 + 0{\cdot}7678W$.

These results bring out clearly the marked difference between summer and winter, which has already been shown in the case of the correlations. But in both cases there is also a difference between the two decades reduced; this is partly due, no doubt, to the fact that the Wilmington observations were taken at different hours in the two decades, but this would not account for the whole difference, which has been found also in dealing with other stations. It may be due to variations corresponding to those of random sampling; or it may indicate a gradual change, whether periodic or progressive, in the physical constants involved; and this question can only be settled by dealing with observations extending over a longer period.

The results of the predictions are shown in Table VI.

The theoretical mean error $= 0{\cdot}7979\sigma\sqrt{1-r^2}$

$$= 0{\cdot}175 \text{ for summer and } 0{\cdot}239 \text{ for winter,}$$

taking the mean summer and mean winter values of σ and r for the two decades. These theoretical errors are slightly larger than the actual mean errors for the forty random dates here considered. We should presumably improve the predictions by taking the interval for which the correlation is a maximum; but even without this improvement the degree of accuracy attained, though not very great, might be

Table VI.—Comparison of Predicted and Observed Heights at Halifax.

A. *Summer Months.*

Date.	Predicted height.	Observed difference.	Date.	Predicted height.	Observed difference.
June 30, 1879 ..	29·738	−0·152	May 8, 1889....	29·844	+0·017
Aug. 17, 1880 ..	29·930	+0·264	April 20, 1890 ..	29·848	+0·245
April 11, 1881 ..	29·789	−0·030	July 1, 1891....	29·771	+0·120
July 4, 1882 ...	29·782	+0·117	Aug. 12, 1892 ..	29·876	−0·061
Sept. 16, 1883 ..	29·922	+0·035	Sept. 7, 1893 ...	29·881	+0·039
April 30, 1884 ..	29·868	−0·306	June 9, 1894 ...	29·902	+0·021
June 14, 1885 ..	29·891	−0·033	July 8, 1895....	29·782	−0·064
March 31, 1886	29·853	+0·235	April 26, 1896 ..	29·842	+0·386
May 16, 1887....	29·991	+0·153	May 13, 1897...	29·875	+0·036
July 20, 1888...	29·924	+0·152	June 17, 1898 ..	29·897	−0·117
Mean error ..	..	0·148	Mean error..	..	0·111

B. *Winter Months.*

Date.	Predicted height.	Observed difference.	Date.	Predicted height.	Observed difference.
Dec. 25, 1879 ...	30·016	−0·396	Nov. 16, 1889...	29·922	+0·194
Oct. 1, 1880	29·905	+0·190	Jan. 5, 1890....	30·126	+0·128
Nov. 25, 1881...	29·502	+0·068	Oct. 28, 1891...	29·683	+0·069
Feb. 21, 1882...	29·951	+0·158	Feb. 14, 1892...	29·888	−0·240
Jan. 11, 1883 ...	29·593	+0·053	Dec. 2, 1893....	29·766	−0·240
Sept. 30, 1884 ..	29·824	+0·102	Jan. 28, 1894...	29·924	+0·146
March 7, 1885 ..	29·894	+0·117	Oct. 1, 1895	29·718	+0·004
Feb. 1, 1886....	29·646	+0·199	Dec. 10, 1896...	29·730	−0·425
Dec. 9, 1887	29·898	+0·395	Feb. 25, 1897...	29·958	+0·211
Nov. 2, 1888....	30·108	−0·004	Nov. 9, 1898 ...	30·114	+0·029
Mean error ..	..	0·168	Mean error..	..	0·169

sufficient to be of service in practical meteorology. Still better results could probably be obtained by making the prediction for one station depend upon observations at two or more other stations, so situated as to give suitable intervals for maximum correlation*; and it may also be desirable to take into account the height observed on the previous day at the station for which the prediction is required, for the results given in Table III show that in some cases, at any rate, there is a

* It would be of little use to predict for intervals of only 5 or 6 hours; at least 24 hours would be desirable, if this does not make it necessarv to take distances so great as seriously to weaken the correlation.

considerable correlation between the barometric readings on successive days at the same station. Moreover, it is possible that the correlation between the daily rise or fall at different stations might give better results for prediction than the correlation between the actual heights observed; but what has already been done is sufficient to indicate the manner in which modern statistical methods may aid the meteorologist in this part of his work. As regards the more fundamental, though less immediately useful, task of meteorological science, the intervals of maximum correlation, a few examples of which have been considered in this paper, will, I believe, be found ultimately to be of importance in relation to the physics of the atmosphere. Although considerable labour is involved in the calculation of these intervals, even for a single pair of stations, yet the increase of theoretical knowledge which would result from an adequate investigation of the manner in which they depend upon local conditions, and upon the positions and distances of the places considered, would probably be amply sufficient to justify the labour expended upon the matter.

8. The conclusions to be drawn from the results given in this paper are as follows:—

(i) The correlation between the barometric readings at two stations upwards of 1000 miles apart depends upon the interval between the readings. In the case of Halifax and Wilmington, the correlation is sensible for at least 9 days, and it reaches a maximum for an interval of about 16 hours in summer and 23 in winter. For these stations, and also for St. Helena and Cape Town, the observation at the more easterly station should be taken later for maximum correlation.

(ii) There is a considerable correlation between the daily rise at Halifax and Wilmington, and this correlation changes with the interval in a manner somewhat analogous to that in which the correlation between simultaneous heights at two stations approximately on the same meridian depends upon the distance between them.

(iii) There are considerable differences between the summer and winter correlations, and these differences are of the same general nature for both pairs of stations considered.

(iv) It is possible to predict the barometric height at one station from an earlier height at a second station more than 1000 miles away, with a fair degree of accuracy, the mean observed error for forty dates, taken at random, for Halifax and Wilmington, being 0″·15.