551·501·45 : 551·52

Statistical study of maximum temperature at Poona with special reference to forecasting of weekly maximum temperature

V. K. RAGHAVENDRA

Meteorological Office, Poona

(*Received 16 December 1953*)

ABSTRACT. Seventytwo years' maximum temperature data of Poona have been studied in this paper. The co-efficients of variability of mean monthly maximum temperatures have maxima in June and October. The frequency distribution of daily maximum temperature during April and May have been studied and found to be skew. The curve $dz/dx = kz^m (1-z)^n$ has been fitted to the frequency distribution. The inter-monthly correlations between the mean maximum temperatures have been found. The winter months are highly correlated. The inter-weekly correlations during the winter months have been calculated. The regression formulae for forecasting of weekly maximum temperature for the week Nos. 46, 47, 48, 49, 50, 51, 52, 1, 2 and 3 of the Grower's year based on the weekly maximum temperatures of the preceding four weeks have been worked out, and comparative diagrams are given. Also the frequency of spells of months with mean maximum temperature above and below the corresponding normal were worked out and the persistency of months of either type has been studied.

1. Introduction

Temperature is an important element of weather which controls the growth and yield of crops directly and also affects indirectly the incidence and intensity of pests and diseases. Forecasts of temperature, especially its extremes in the form of heat and cold waves, would, therefore, be of practical value to agriculturists. Forecast of maximum temperature during winter in Deccan will be of utility for plant physiologists in regard to the attack of pests like Cumin Mildew. An approach to this problem, using maximum temperature data of Poona for 52 years from 1880 to 1931 was made by Kalamkar (1934). The investigation was also extended by Kalamkar (1936) to a few selected stations to discuss the relation, if any, between the periods of high inter-monthly correlation co-efficients of mean monthly maximum temperature and seasonal changes in weather.

In his analysis of maximum temperature data at Poona from 1880 to 1931 Kalamkar found that the co-efficient of variability of the mean maximum temperature is comparatively higher for June and October, these being months in which the monsoon sets in and withdraws respectively. The inter-monthly correlation co-efficients were worked out. After finding that the autumn and winter months were highly correlated, regression formulae for forecasting weekly maximum temperatures in December based on the preceding four weeks' temperatures were worked out by him.

Kalamkar's conclusions and formulae were based on data relating to 1880 to 1931. Now that more than 20 years' data have accumulated, one would like to know how the additional data modifies the conclusions.

Herein we shall study the data for the period 1880 to 1951 and the inter-monthly and inter-weekly correlations with a view to forecasting the maximum temperatures during the standard week Nos. 46, 47, 48, 49, 50, 51, 52, 1, 2 and 3. We shall also be studying the frequency distribution of the daily maximum temperature at Poona for the two months April and May which are the hottest months in Poona, using the data for the period 1901 to 1950 and the persistency of spells of months having their mean maximum temperatures above and below the corresponding normal.

TABLE 1

Mean monthly maximum temperature and standard deviation

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Means (°F)	86·5	90·6	97·1	100·8	98·8	89·3	82·4	81·9	84·6	89·5	86·9	85·1
S. D. (°F)	1·59	1·75	1·96	2·15	1·80	2·55	1·66	1·45	1·90	2·65	2·19	1·92

2. Summary of 72 years' data of maximum temperature, its seasonal behaviour and variability

The average maximum temperatures for different months for the period of 72 years were tabulated and the mean and the standard deviation were worked out. At Poona, the site of temperature observations was changed about 1930. The corrections accepted by the India Meteorological Department were applied to the data prior to the change of site. The results are set out in Table 1. The homogeneity of the series has been tested and it was found that other corrections applied for change of site were quite adequate to render the series homogeneous.

The two maxima, one in April and the other in October, which is a well known feature over large areas of India, are brought out clearly in the table above.

From Table 1 we see that the mean maximum temperature for April and May are 100·8° and 98·8°F respectively and their standard deviations are only 2·15° and 1·80°F respectively. As the highest maximum temperature recorded at Poona is 110°F which does not come within even five times the standard deviation, an attempt was made to study the frequency distribution of the daily maximum temperature during these two months. For this investigation the data relating to the fifty-year period, 1901 to 1950, were considered. Table 2 gives the distribution of frequency of daily maximum temperatures during this period.

From Table 2, the mean, standard deviation and the Pearsonian constants, β_1, β_2, κ, sk and d with their probable error (p.e.)

TABLE 2

Frequency of daily maximum temperature during 1901—50

Temperature (°F)	Frequency April	Frequency May
<80	1	1
81	1	0
82	1	0
83	1	2
84	0	1
85	3	8
86	0	3
87	2	4
88	3	2
89	4	5
90	8	3
91	18	13
92	16	22
93	31	37
94	26	59
95	44	88
96	53	104
97	98	140
98	94	132
99	141	195
100	147	148
101	172	143
102	149	105
103	162	108
104	136	69
105	100	81
106	46	32
107	30	27
108	11	12
109	2	5
110	0	1

TABLE 3

Month	Mean (°F)	Standard deviation (°F)	β_1	β_2	κ	*sk*	*d* (°F)
April values	100·4	3·91	0·83	4·72	0·80	0·37	1·44
p. e. of the above	0·07	0·05	0·17	0·71	0·67	0·05	0·18
May values	99·3	4·04	0·13	4·27	0·05	0·11	0·45
p. e. of the above	0·07	0·05	0·07	0·85	0·03	0·03	0·12

were calculated for the distributions and are given in Table 3.

It is seen from Table 3 that the distributions are not normal and are definitely skew. An attempt was made to fit curves to the distributions. The main consideration for deciding the form of curve was to obtain similar curves for both the distributions. Hence the curve $dz/dx = kz^m(1-z)^n$ was fitted as suggested by Krechewsky (*see reference*) where z is the area of the curve upto the ordinate at x; k, m and n are constants to be determined from the distribution. The curve gives a sufficiently good fit to the data and the values of k, m and n are given in Table 4.

TABLE 4

Month	*k*	*m*	*n*
April	0·3438	0·9664	0·6744
May	0·3521	0·8664	0·8590

The curve fitted for April is shown along with the actual frequencies in Fig. 1. From the figure it is seen that there is a close agreement between the curve fitted in and the actual frequencies. Curves of these types are fitted in to data relating to other months as well and shall be discussed separately.

Next, the mean monthly maximum temperatures were plotted to detect the presence of any 'secular trend'. The diagrams, in general, showed an absence of any 'secular trend', thereby confirming Kalamkar's findings.*

3. Inter-monthly correlations and their significance

The inter-monthly correlations between the mean maximum temperatures were calculated using the data for the longer period of 72 years. They are set out in Table 5.

From Fisher's table for significance of correlation co-efficients, the value of correlation co-efficient for 5 per cent level of significance is 0·30. Out of the 76 correlation co-efficients in the above table, 22 exceed the 5 per cent value. Almost all the correlation co-efficients which were found to be significant by Kalamkar (1934) are found to be significant here also although generally the values of correlation co-efficients are somewhat lower. The correlation co-efficients between the winter months being very high is also established.

4. Inter-weekly correlations and regression formulae

In view of the high correlation co-efficient between the winter months' maximum temperature, the investigation was extended to the smaller period of a week and formulae for forecasting the weekly maximum temperature in winter based on those of the preceding weeks was derived as suggested by Kalamkar (1934).

A slight change was made in the seven day periods chosen by Kalamkar. Instead of the periods chosen by him, the standard weeks of the Grower's year (Ramdas 1942) adopted for crop weather work in India were considered. The serial number of the weeks

*The absence of trend in Poona maximum temperature has also been indicated by Pramanik and Jagannathan (1954)

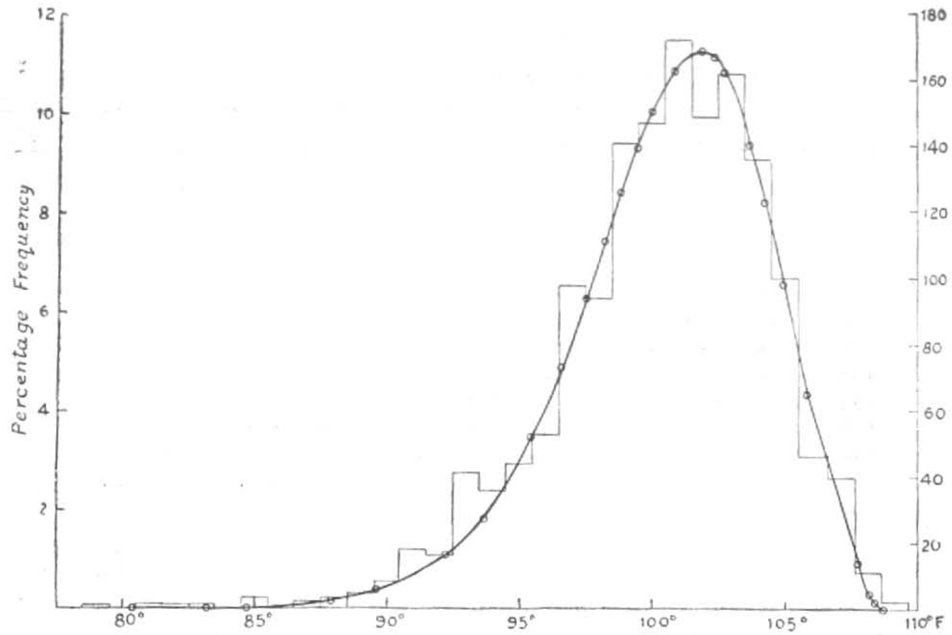

Fig. 1. Frequency distribution of daily maximum temperatures at Poona during April showing the actual and the curve fitted

considered, their dates and mean maximum temperatures (total of the daily values for the week), their standard deviations and co-efficients of variation, are given in Table 6.

The weekly maximum temperatures for different weeks were also plotted against the year. The diagrams showed in general the absence of any 'secular trend'. The inter-weekly correlations have been worked out and are given in Table 7. Out of the 67 correlations, 63 are significant at 1 per cent level and they are all positive. From an examination of this table we see that the correlation co-efficients of any one week with the succeeding weeks decrease. But after about 3 weeks in some cases, it shows a tendency to increase. Hence, in order to cover this tendency, the regression formulae for forecasting any one week's maximum temperature was based on the preceding four weeks.

With the help of Tables 6 and 7 the regressions of the mean maximum temperature during the weeks 46 to 52 and 1 to 3 depending on the four preceding weeks were worked out, denoting M_r as accumulated total maximum temperature for week No. r. The regression formulae are given below—

$$
\begin{aligned}
M_{46} &= 150 + 0{\cdot}1312\, M_{42} + 0{\cdot}1361\, M_{43} + 0{\cdot}1001\, M_{44} + 0{\cdot}3729\, M_{45} \\
M_{47} &= 238 + 0{\cdot}0408\, M_{43} + 0{\cdot}1025\, M_{44} - 0{\cdot}1439\, M_{45} + 0{\cdot}5998\, M_{46} \\
M_{48} &= 203 - 0{\cdot}0421\, M_{44} + 0{\cdot}1539\, M_{45} + 0{\cdot}2474\, M_{46} + 0{\cdot}2927\, M_{47} \\
M_{49} &= 205 + 0{\cdot}0907\, M_{45} + 0{\cdot}0925\, M_{46} + 0{\cdot}0505\, M_{47} + 0{\cdot}4185\, M_{48} \\
M_{50} &= 134 + 0{\cdot}1926\, M_{46} + 0{\cdot}1521\, M_{47} - 0{\cdot}1165\, M_{48} + 0{\cdot}5375\, M_{49} \\
M_{51} &= 227 - 0{\cdot}0127\, M_{47} + 0{\cdot}1409\, M_{48} + 0{\cdot}3667\, M_{49} + 0{\cdot}1208\, M_{50} \\
M_{52} &= 181 + 0{\cdot}0713\, M_{48} - 0{\cdot}0709\, M_{49} + 0{\cdot}2734\, M_{50} + 0{\cdot}5648\, M_{51} \\
M_{1} &= 265 + 0{\cdot}2834\, M_{49} - 0{\cdot}0003\, M_{50} + 0{\cdot}0892\, M_{51} + 0{\cdot}1619\, M_{52} \\
M_{2} &= 328 - 0{\cdot}1168\, M_{50} + 0{\cdot}1601\, M_{51} + 0{\cdot}0599\, M_{52} + 0{\cdot}3463\, M_{1} \\
M_{3} &= 256 - 0{\cdot}1413\, M_{51} + 0{\cdot}1543\, M_{52} - 0{\cdot}0413\, M_{1} + 0{\cdot}5902\, M_{2}
\end{aligned}
$$

TABLE 5

Correlations between mean monthly maximum temperatures

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Feb	0·20											
Mar	0·45	0·32										
Apr	0·17	0·27	0·31									
May	0·13	0·22	0·22	0·30								
Jun	–0·13	0·16	–0·08	0·06	0·47							
Jul	0·01	0·19	0·07	0·07	0·21	0·32						
Aug	–0·04	0·01	0·17	–0·06	–0·03	–0·08	0·16					
Sep	–0·01	0·03	0·05	–0·01	0·10	0·06	0·19	0·01				
Oct	–0·09	0·02	0·15	–0·04	0·06	0·05	0·03	–0·07	0·46			
Nov	–0·06	0·00	0·06	–0·18	0·04	0·12	0·11	0·15	0·41	0·62		
Dec	0·15	–0·04	0·12	–0·04	0·10	0·04	0·02	0·10	0·37	0·54	0·64	
Next year Jan									0·29	0·34	0·52	0·42
Next year Feb										0·24	0·38	0·28
Next year Mar											0·64	0·28
Next year Apr												0·08

TABLE 6

Mean of totals of weekly maximum temperatures for standard weeks 42 to 3

Week No.	Date	Total mean (°F)	S. D. (°F)	Co-efficient of variation (%)
42	15—21 Oct	628·1	21·9	3·49
43	22—28 Oct	626·4	22·1	3·53
44	29 Oct—4 Nov	621·9	19·0	3·06
45	5—11 Nov	613·0	20·4	3·33
46	12—18 Nov	608·9	18·7	3·07
47	19—25 Nov	604·8	23·3	3·84
48	26 Nov—2 Dec	598·9	19·1	3·19
49	3—9 Dec	597·4	17·2	2·87
50	10—16 Dec	595·0	18·4	3·09
51	17—23 Dec	595·3	16·9	2·84
52	24—31 Dec (8 days)	680·5	20·7	3·04
1	1—7 Jan	597·2	17·6	2·95
2	8—14 Jan	601·4	14·8	2·46
3	15—21 Jan	607·4	16·6	2·73

TABLE 7

Correlations between weekly maximum temperatures

Week No.	42	43	44	45	46	47	48	49	50	51	52	1	2
43	0·7168												
44	0·4237	0·7047											
45	0·2875	0·4117	0·5269										
46	0·4290	0·5102	0·4946	0·5706									
47	0·2547	0·2922	0·2835	0·2094	0·4719								
48	0·2653	0·3404	0·2655	0·3550	0·4833	0·4930							
49	0·4894	0·3460	0·2621	0·3453	0·4200	0·3685	0·5871						
50	0·3652	0·3402	0·2849	0·3489	0·4397	0·4108	0·3636	0·5847					
51	0·4121	0·3498	0·1591	0·3362	0·4051	0·2445	0·4100	0·5249	0·3811				
52	0·3441	0·2984	0·2626	0·3461	0·3274	0·3005	0·3091	0·3644	0·4084	0·5507			
1								0·3903	0·2715	0·3353	0·3378		
2									0·0709	0·3122	0·2647	0·4625	
3										0·1116	0·2379	0·2166	0·5127

TABLE 8

Analysis of Variance

Week No.	Due to	D.F.	S. S.	M. S.	$\frac{1}{2} \log_e 1/100$ (M.S.)
46	Regression	4	10898·05	2724·51	1·6525
	Deviation from regression	67	14340·61	214·04	0·3805
	Total	71	25238·66	355·47	Z = 1·2720
47	Regression	4	5722·92	1430·73	1·3303
	Deviation from regression	66	18475·48	279·93	0·5147
	Total	70	24198·40	345·69	Z = 0·8156
48	Regression	4	8947·04	2236·76	1·5539
	Deviation from regression	67	17383·27	259·45	0·4769
	Total	71	26330·31	370·85	Z = 1·0770
49	Regression	4	8029·27	2007·32	1·4997
	Deviation from regression	67	13178·51	196·69	0·3384
	Total	71	21207·78	298·70	Z = 1·1613
50	Regression	4	10080·64	2520·16	1·6134
	Deviation from regression	67	14198·35	211·92	0·3754
	Total	71	24278·99	341·96	Z = 1·2380
51	Regression	4	6082·95	1520·74	1·3608
	Deviation from regression	66	14240·97	215·77	0·3844
	Total	70	20323·92	290·34	Z = 0·9764
52	Regression	4	10849·95	2712·49	1·6503
	Deviation from regression	67	19930·04	297·46	0·5451
	Total	71	30779·99	433·52	Z = 1·0052
1	Regression	4	4483·73	1120·93	1·2084
	Deviation from regression	67	17856·71	266·52	0·4901
	Total	71	22340·44	314·65	Z = 0·7183
2	Regression	4	4408·11	1102·03	1·2001
	Deviation from regression	67	11346·39	169·35	0·2635
	Total	71	15754·50	221·89	Z = 0·9366
3	Regression	4	5653·27	1413·32	1·3242
	Deviation from regression	66	13807·21	209·20	0·3690
	Total	70	19460·48	278·01	Z = 0·9552

Significance of the regression formula was tested by the z test of Fisher (1944) by analysing the total sum of squares of deviations. Also the significance of the multiple correlation co-efficient was tested by Wishart's (1928) tables. The results of analysis of variances are shown in Table 8.

The value of z obtained exceed even the one per cent point and are thus highly significant.

The values of the multiple correlation coefficients in forecasting mean weekly maximum temperatures for the weeks 46 to 52 and 1 to 3 are 0·657, 0·486, 0·573, 0·615, 0·644, 0·547, 0·594, 0·448, 0·529 and 0·539 respectively. These correlation co-efficients are all significant and this suggests the possibility of forecasting the weekly maximum temperatures considered. The forecasted values from the regression equations and the actuals of 10 weeks maximum temperatures are shown in Figs. 2 to 11.

The addition of 20 years' data, therefore, has not altered any of the conclusions drawn by Kalamkar (1934) earlier and have confirmed the results obtained by him. The regression forecasts of weekly maximum temperatures during winter has been proved to be possible by the highly significant multiple correlation co-efficients and the work was extended to forecast the maximum temperatures during November to January.

5. Persistence of monthly maximum temperature above or below normal

The high and significant inter-monthly correlations suggested that a study can be made of the persistency in months with mean maximum temperature, above and below their corresponding normals. For this, the departures of the monthly mean maximum temperatures from their average for the 72 years (1880–1951) were tabulated. The correlation between any two successive values (*i.e.*, the serial correlation) was found to be 0·3765 for 863 pairs of values considered with 861 degrees of freedom. This is found to be significant.

TABLE 9

Current month	Preceding month		Total
	A	B	
A	267	171	438
B	171	254	425
Total	438	425	863
Fraction of A to total	0·6096	0·4024	
Fraction of B to total	0·3904	0·5976	
Standard error	0·0239	0·0242	

The months were classified as A and B according as their mean maximum temperature was above or below the normal. Out of 864 (72 × 12) months 439 were A and 425 were B, thus bringing the probability of any month being A as 0·5081 with standard error 0·01701. The probability is not significantly different from 0·5 (which is the random probability). If there be any tendency for the A months occurring together consecutively then the probability of A will be greater if the previous month was A than if the previous month was B. To test this the A months were classified according as the preceding month was A or B (Table 9).

The total number of months included is 863, the first was omitted as it was not known whether A or B preceded it. The χ^2 for the above '2×2' contingency table is 36·24 which is highly significant. This suggests the absence of independence. The chance of A being followed by A is 0·6096 and B being followed by B is 0·5976. Both these values are significantly different from their random probabilities of occurrence, in other words, there is a tendency for A and B months to persist. Also the proportion of cases where the succeeding month was different is 0·3963 and where it was the same is 0·6937. These also show a tendency for persistence of A and B months.

To study the spells in greater detail frequencies of spells of A and B of different lengths were worked out (Table 10).

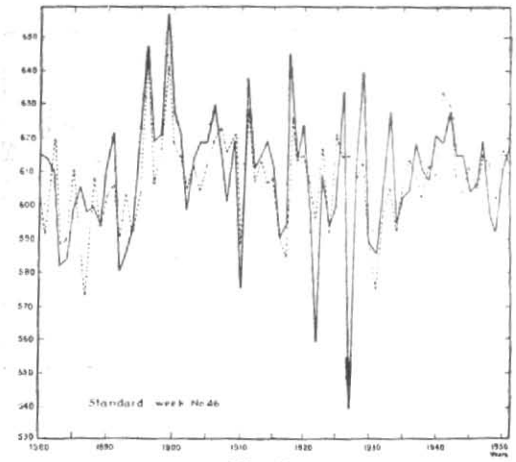

Fig. 2

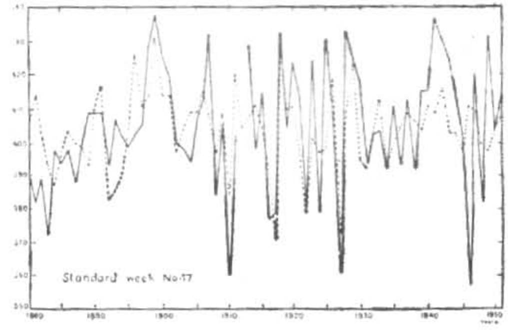

Fig. 3

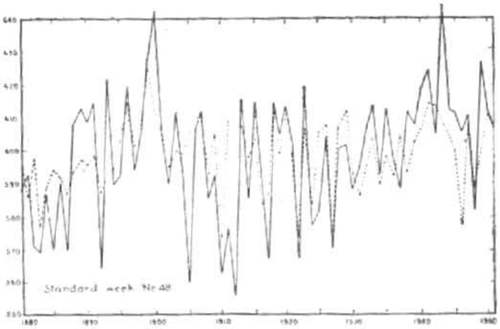

Fig. 4

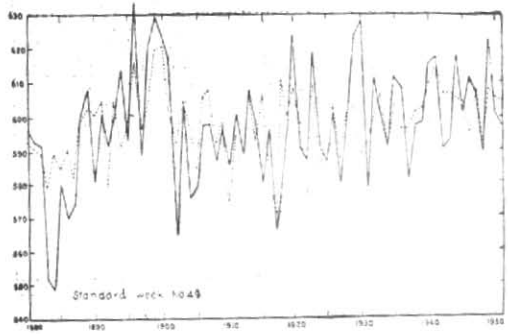

Fig. 5

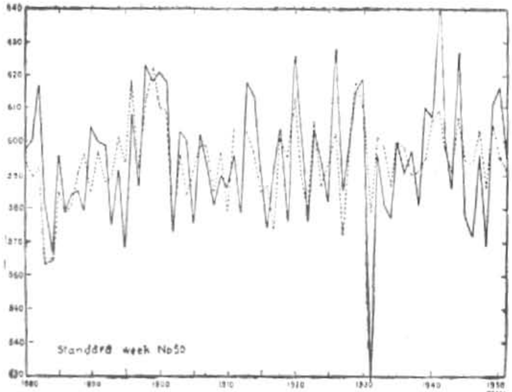

Fig. 6

Figs. 2-6. Calculated and observed weekly maximum temperature (total) for the standard week Nos. 46-50

(— observed value ; .... regression value)

The ordinate in the above diagrams refer to weekly totals of maximum temperature in °F

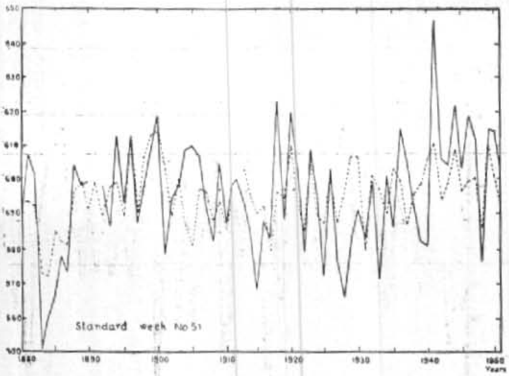

Fig. 7

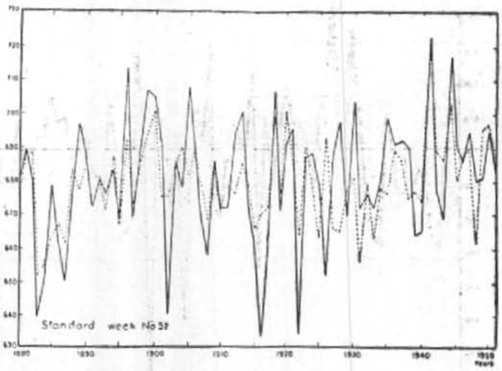

Fig. 8

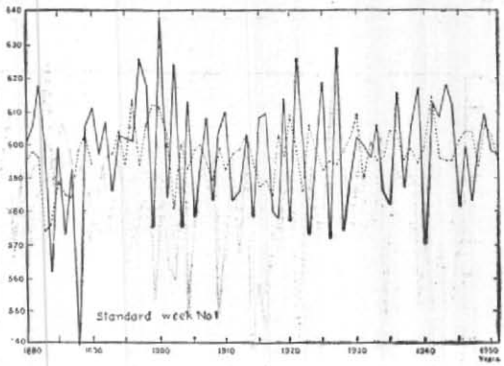

Fig. 9

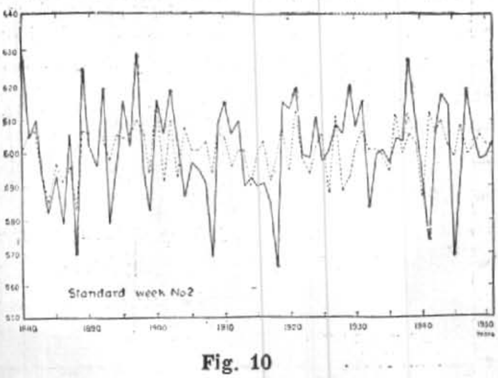

Fig. 10

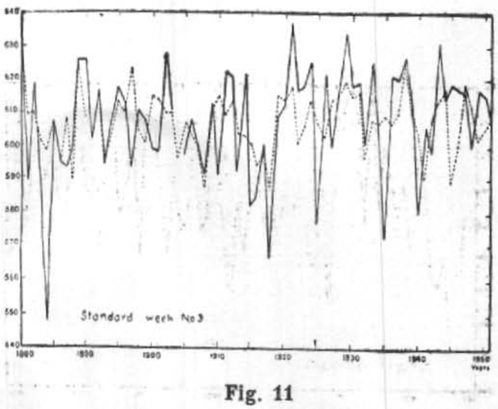

Fig. 11

Figs. 7-11. Calculated and observed weekly maximum temperature (total) for the standard week Nos. 51, 52, 1-3

(— observed value ; .... regression value)

The ordinate in the above diagrams refer to weekly totals of maximum temperature in °F.

TABLE 10

	Duration in months												
	1	2	3	4	5	6	7	8	9	10	11	12	13
A	82	35	18	16	5	3	3	4	..	..	1	3	2
B	73	38	30	8	9	3	3	1	3	..	2	..	1

TABLE 11

	Duration in months (r)							
	1	2	3	4	5	6	7	8
A	52	51	67	57	76	81	77	60
B	57	61	50	73	59	77	70	86

By summing the sequences from the longest to the shortest we can compute the chance (*i.e.*, the percentage frequency) of recurrence of A or B month after a run of r months of the same type. These percentage frequencies are given upto r = 8 (Table 11).

The table shows that the chance of occurrence of A month after a run of 4 is as high as 75 per cent (compared to the 51 per cent chance of occurrence of A month at random) and that it can occur with such a chance for 3 consecutive months. In the case of B months the tendency to persist starts after 5 months of the same type. A knowledge of this persistency would be of value in the forecasting of maximum temperature.

6. Acknowledgement

The author is grateful to Dr. L. A. Ramdas for his valuable guidance in the preparation of this note and to Messrs S. P. Venkiteshwaran, V. K. Ramabhadran and P. Jagannathan for going through the paper and for helpful suggestions.

REFERENCES

Cochran, W. G.	1938	*Quart. J. R. met. Soc.*, **164** pp. 631–34.
Fisher, R. A.	1944	*Statistical Methods for Research Workers.*
Fisher and Yates	..	*Table for Statisticians and Biometricians.*
Kalamkar, R. J.	1934	*India met. Dep. Sci. Notes*, **5**, 59.
	1936	*ibid*, **7**, 70.
Krechewsky, S.	..	*A Method of Curve Fitting*, Paper 8, Physical Dep., Min. of Public Works, Egypt.
Pramanik, S. K. and Jagannathan, P.	1954	*Indian J. Met. Geophys.*, **5**, 1, p. 29.
Ramdas, L. A.	1942	*Indian Farming*, **3**, 5.
Wishart, J.	1928	*Quart, J. R. met. Soc.*, **54**, pp. 258–59.