INFLUENCE OF HUMIDITY AND TEMPERATURE ON THE YIELD OF COTTON

By K. RAGHAVAN NAIR AND PURNENDU BOSE
Statistical Laboratory, Calcutta.

Fisher (1924) devised a special technique by which he studied the influence of weekly rainfall on the annual yield of wheat at Rothamsted. This paper is a generalisation of the above method when the influence of more than one meteorological factor is to be studied.

General formulae. Let $M_1, M_2, \ldots, M_k$ be k meteorological factors which influence the yield of the crop y. Let q be the subdivisions of the year for each of which records of the meteorological factors are available. The multiple linear regression equation of the yield of the crop upon the k meteorological factors is written as

$$Y = C + \sum_{s=1}^{q} a_{1s} m_{1s} + \sum_{s=1}^{q} a_{2s} m_{2s} + \ldots + \sum_{s=1}^{q} a_{ks} m_{ks} \quad \ldots \quad (1)$$

where $m_{11}, m_{12}, \ldots, m_{1q}, m_{21}, m_{22}, \ldots, m_{2q}, m_{k1}, m_{k2}, \ldots, m_{kq}$ are the measurements of the meteorological factors in the different intervals of time. If the subdivisions of time were made infinitely small we should replace the linear regression function by a regression integral of the form

$$Y = C + \int_0^T a_1 m_1 dt + \int_0^T a_2 m_2 dt + \ldots + \int_0^T a_k m_k dt \quad \ldots \quad (2)$$

where $m_i dt$ in general is the effect of the meteorological factor in the element of time dt. The integral is taken over the whole period.

If $T_0, T_1, T_2, \ldots, T_n$ be a series of orthogonal functions of time such that

$$\int_0^T T_r T_s dt = 0 \quad (r \neq s), \qquad \int_0^T T_r^2 dt = 1.$$

then we may represent the series of values of each of the meteorological factors as a time series using the following orthogonal polynomials.

$$\left.\begin{aligned} m_1 &= \rho_{10} T_0 + \rho_{11} T_1 + \rho_{12} T_2 + \ldots + \rho_{1n} T_n \\ m_2 &= \rho_{20} T_0 + \rho_{21} T_1 + \rho_{22} T_2 + \ldots + \rho_{2n} T_n \\ m_k &= \rho_{k0} T_0 + \rho_{k1} T_1 + \rho_{k2} T_2 + \ldots + \rho_{kn} T_n \end{aligned}\right\} \quad \ldots \quad (3)$$

where $$\rho_{is} = \int_0^T m_i T_s dt$$

The regression values $a_{11}, a_{12}, \ldots, a_{1q}$, for the meteorological factors in (1) may be expected to lie on a continuous time curve so that we may express the regression function in (2) in the form

$$\left.\begin{aligned} a_1 &= \alpha_{10} T_0 + \alpha_{11} T_1 + \alpha_{12} T_2 + \ldots \\ a_2 &= \alpha_{20} T_0 + \alpha_{21} T_1 + \alpha_{22} T_2 + \ldots \\ a_k &= \alpha_{k0} T_0 + \alpha_{k1} T_1 + \alpha_{k2} T_2 + \ldots \end{aligned}\right\} \quad (4)$$

where $$\alpha_{is} = \int_0^T a_i T_s dt$$

The degree of the polynomials in (3) and (4) should be the same and in practice seldom exceeds 4 or 5. Now using the relations (3) and (4) the relation (2) stands as

$$\begin{aligned} Y = C &+ (\rho_{10}\alpha_{10} + \rho_{11}\alpha_{11} + \ldots + \rho_{1n}\alpha_{1n}) + (\rho_{20}\alpha_{20} + \rho_{21}\alpha_{21} + \ldots + \rho_{2n}\alpha_{2n}) \\ &+ \ldots + (\rho_{k0}\alpha_{k0} + \rho_{k1}\alpha_{k1} + \ldots + \rho_{kn}\alpha_{kn}) \quad \ldots \quad (5) \end{aligned}$$

Now the values $\rho_{10}, \rho_{11}, \ldots, \rho_{k0}, \rho_{k1}, \ldots$ may be obtained for each year by fitting orthogonal polynomials for each meteorological factor, then correlating this series with the yield of the crop for several years, we can get value of α_{is} ($i = 1$ to k; $s = 1$ to n) as partial regression coefficients using relation (5).

By substituting these values in (4) we get values of $a_1, a_2, \ldots, a_k$ which will give us the effect of fluctuations in each meteorological factor at any point of time in the season on the yield of the crop at the end of the season. In this paper one actual example has been worked out with two meteorological factors. In this special case equation (5) will reduce to

$$Y=C+(\rho_{10}\alpha_0+\rho_{11}\alpha_1+\ \ldots\ +\rho_{1s}\alpha_s)+(\rho_{20}\beta_0+\rho_{21}\beta_1+\ \ldots\ +\rho_{2s}\beta_s) \qquad \ldots \quad (5\cdot1)$$

The description and the method of analysis of the experiment is discussed in the next paragraph.

Numerical Illustration. The application relates to yield of cotton recorded in the agricultural research station, Sarkand (Sind) and supplied to us for investigation by the Director of Agriculture, Sind. In certain years there were low yields in particular tracts; and it was apprehended that such low yields were due to climatic factors. Yield figures for 1931-1941 (year 1937 was omitted) along with certain humidity and maximum temperature figures for four months (August to November) for each year were sent to us for analysis. The number of years is too small, but the material has been analysed to elucidate the practical procedure. For each day there were four readings for temperature and humidity from 10 A.M. to 4 P.M. at inter- vals of 2 hours; the mean of these four readings was taken as the representative figure for the day. We have thus 119 individual figures for temperature and humidity for each year; these were analysed by calculating the coefficients of the polynomial up to 3rd degree. The distribution of temperature and humidity in each season is thus represented by four figures.

The computation of temperature and humidity coefficients involves a great deal of labour. The method of successive summation first given by Hardy and subsequently by Fisher was used. The method may be briefly described as follows.

Let the whole interval be divided into two parts (1 to m) and (m to n). Then,

$$s_1=\sum_{i=1}^{m} x_i,\ s_2=\sum_{i=1}^{m}\left\{(m+1-i)\,x_i\right\},\ s_3=\sum_{i=1}^{m}\left\{(m+1-i)\,(m+2-i)\,x_i/2!\right\}$$

$$s_k=\sum_{i=1}^{m}\left\{(m+1-i)\,(m+2-i)\ldots(m+\overline{k-1}-i)x_i/(k-1)!\right\}$$

where x_i are the numbers. The remaining n numbers are to be summed backwards from the bottom of the column, dropping one term at the end of each summation and so giving

$$s'_1=\sum_{i=m}^{n} x_i,\ s'_2=\sum_{i=m}^{n}\left\{(m+1-i)\,x_i\right\},\ s'_3=\sum_{i=m}^{n}\left\{(m+1-i)(m+2-i)x_i/2!\right\}$$

$$s'_k=\sum_{i=m}^{n}\left\{(m+1-i)\,(m+2-i)\ \ldots\ (m+k-1-i)\,x_i/(k-1)!\right\}$$

Taking alternately sums and differences we obtain

$$S'_1=s_1+s'_1,\quad S'_2=s_2-s'_2,\quad S'_3=s_3+s'_3,\quad S'_4=s_4-s'_4$$

and so on, from which the final sums of the whole column may be obtained from the equation.

$$S_1=S'_1,\quad S_2=S'_2+MS'_1,\quad S_3=S'_3+MS'_2+M\,(M+1)\,S'_1/2!$$

$$S_4=S'_4+MS'_3+M\,(M+1)\,S'_2/2!+M(M+1)\,(M+2)\,S'_1/3! \quad \text{where} \quad M=n-m.$$

The series $S_1, S_2, S_3, S_4, \ldots$ has to be divided by a series of numbers n, $n\,(n+1)/2!$, $n(n+1)\,(n+2)/3!$.. $n(n+1)(n+2)(n+3)/4!$ etc. yielding a series of numbers.

$$a=\frac{1}{n}\Sigma(x_i),\quad b=\frac{2!}{n(n+1)}\Sigma\{(n+1-i)x_i\},\quad c=\frac{3!}{n(n+1)(n+2)}\Sigma\{(n+1-i)\,(n+2-i)x_i/2!\}$$

$$d=\frac{4!}{n(n+1)(n+2)(n+3)}\,\Sigma\{(n+1-i)\,(n+2-i)\,(n+3-i)\,x_i/3!\}$$

In the present case as we have fitted a third degree curve, it is sufficient to calculate up to c. From the values of $a, b, c, d, \ldots$ we can finally calculate ρ-values from the following relations

$$\rho_0=\sqrt{n}\,.a,\quad \rho_1=(a-b)\sqrt{3n\,(n+1)/(n-1)},\quad \rho_2=(a-3b+2c)\sqrt{5n(n+1)\,(n+2)/(n-1)(n-2)}$$

$$\rho_3=(a-6b+10c-5d)\sqrt{7n(n+1)\,(n+2)\,(n+3)/(n-1)(n-2)(n-3)}$$

Polynomial values of Meteorological Factors. In table 1, columns (4) and (5) gives the values of S_r for different values of r, and the last two columns give the values of the constants a, b, c, and d, for each year.

TABLE 1. WEIGHTED MEANS OF THE SERIES OF READINGS FOR HUMIDITY AND TEMPERATURE.

year	constant		S_r humidity	S_r temperature	values of the constant humidity	values of the constant temperature
(1)	(2)	(3)	(4)	(5)	(5)	(7)
1931	a	1	6347·4	10569·8	53·0202	88·8218
	b	2	425440·7	649440·8	59·5883	90·9581
	c	3	17665140·7	26517915·2	61·3410	92·0843
	d	4	544473478·2	814500515·4	61·9890	92·7249
1932	a	1	5540·7	10581·2	46·5605	88·9176
	b	2	378878·9	656709·8	53·0040	91·9846
	c	3	16332976·5	26780306·0	56·7157	92·9936
	d	4	517979374·9	819840926·8	58·9726	93·3399
1933	a	1	6000·5	10650·47	50·9286	89·4097
	b	2	390111·7	654479·41	54·6375	91·6038
	c	3	16403590·4	26514585·22	56·9609	92·0709
	d	3	512533025·8	800598382·30	58·3525	92·1738
1934	a	1	5138·7	10006·0	43·1824	84·0840
	b	2	346501·4	632339·4	48·5206	88·5629
	c	3	14742044·0	26017704·8	51·1012	90·3453
	d	4	460845528·0	800895752·2	52·4678	91·1830
1935	a	1	4565·7	10289·1	38·3072	86·4630
	b	2	287612·0	643466·9	40·2818	80·1214
	c	3	11084684·4	25402355·8	41·0104	91·6812
	d	4	371828888·6	811002273·8	42·3328	92·4301
1936	a	1	4122·1	10518·0	34·6305	88·3866
	b	2	257847·0	657457·8	36·1130	92·0869
	c	3	10756772·3	26882440·4	37·3525	93·3483
	d	4	334795700·9	824728218··6	38·1169	93·8963
1938	a	1	4446·4	11107·6	37·3647	93·3412
	b	2	306746·6	692000·4	42·9045	96·9188
	c	3	13107181·0	28179982·2	45·8267	97·8540
	d	4	416306927·8	862326901·0	47·2809	98·1770
1939	a	1	3816·0	11277·6	32·0672	94·7697
	b	2	259626·0	697233·2	36·3623	97·6517
	c	3	11129474·7	28351035·0	38·6467	98·4479
	d	4	350337184·4	867277874·0	39·8803	98·7407
1940	a	1	4167·5	11340·7	35·0210	95·3000
	b	2	270030·4	695120·2	39·0709	97·3506
	c	3	12002542·0	28108950·5	41·6784	97·8157
	d	4	370086284·4	861206903·6	43·1479	98·0607

Table 2 gives the orthogonal polynomial coefficients for humidity and temperature for each year in columns (3) to (8). Col. (2) gives the yield of cotton for each year.

Regression Equations. Having obtained the ρ values for temperature and humidity we have fitted separately (1) yield against humidity; (2) yield against temperature; and (3) yield jointly against humidity and temperature. The respective equations are as follows

$$Y = 112 \cdot 8274 + 2 \cdot 1910\,\rho_{10} - 0 \cdot 7752\,\rho_{11} + 4 \cdot 1904\,\rho_{12} \quad \text{.. (i)}$$

$$Y = 1013 \cdot 4737 + 0 \cdot 0002\,\rho_{20} - 0 \cdot 0122\,\rho_{21} + 16 \cdot 0435\,\rho_{22} \quad \text{.. (ii)}$$

$$Y = 403 \cdot 0320 + 2 \cdot 1283\,\rho_{10} - 0 \cdot 0702\,\rho_{11} + 3 \cdot 5690\,\rho_{12} - 0 \cdot 001584\,\rho_{20} - 0 \cdot 0050\,\rho_{21} + 11 \cdot 4486\,\rho_{22} \quad \text{..(iii)}$$

Values of the multiple correlation coefficients are given below.

$R_1 = 0 \cdot 3079$, $R_2 = 0 \cdot 5099$, $R_3 = 0 \cdot 5646$

The degrees of freedom for testing R_3, R_2, and R_1 are 3, 5 and 2 respectively, but none of them are significant.

TABLE 2. ORTHOGONAL POLYNOMIAL COEFFICIENTS FOR HUMIDITY AND TEMPERATURE

year	yield	humidity			temperature		
		p_{10}	p_{11}	p_{12}	p_{20}	p_{21}	p_{22}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1931	1694·2	600·27	−87·20	−26·56	969·05	−40·70	−2·81
1932	591·5	507·98	−123·89	20·01	970·09	−58·43	−26·25
1933	1701·1	535·03	−70·65	23·47	976·44	−41·23	−33·77
1934	1328·0	471·12	−101·88	−0·65	917·36	−85·32	−22·86
1935	1414·0	416·59	−36·47	18·88	943·31	−60·69	−13·46
1936	1380·0	377·91	−28·06	25·11	964·30	−70·38	−20·01
1938	655·0	407·60	−106·68	3·50	1016·35	−68·15	−42·71
1939	1374·0	349·85	−81·82	6·65	1033·04	−54·90	−32·27
1940	857·0	382·07	−77·34	28·52	1039·72	−30·20	−28·48

Conclusion. On available evidence there is no significant influence of temperature and humidity on the yield of cotton. It must however be remembered that owing to the scantiness of the material analysed here, any but a large influence might well remain undetected.

REFERENCES

1. FISHER R. A. (1924) The influence of rainfall on the yield of wheat at Rothamsted. *Philosophical Transactions of the Royal Society of London*, B, 213, 89-142.

Paper received : 20 March, 1943.