15 1923

ON ERRORS OF OBSERVATION AND UPPER AIR RELATIONSHIPS.

BY

P. C. Mahalanobis, M.A. (Cantab.),

Professor of Physics, Presidency College, Calcutta.

(Received April, 1922.)

Dr. E. H. Chapman in his paper on the "Relationship between pressure and temperature at the same level in the free atmosphere" (1) has discussed the correction for errors of observation in the case of Dines's correlation coefficients between pressure and temperature at the same level in the free atmosphere (2). According to Shaw and Dines the fluctuations of pressure are the real causes of temperature fluctuations in the free atmosphere. This view has far reaching consequences and is based to a great extent on the statistical evidence furnished by high values of the correlation coefficient between pressure and temperature. Chapman believes that he has succeeded in "correcting" Dines's values in such a way as to actually give +1·00 in several instances. A correlation of 1·00 establishes an absolute causal nexus and leaves no room for effects due to other factors such as the region from which the air is brought, whether tropical or arctic. A result of such importance demands close scrutiny. In the present paper I have investigated in detail one important assumption upon which the validity of Dr. Chapman's work appears to depend, and unless I am mistaken the conclusion must be that this assumption is not justified.

2. I shall adopt the notation used by Chapman. $x_1, x_2, x_3, \ldots\ldots$ are "true" departures from a mean value while $a_1, a_2, a_3, \ldots\ldots$ are the corresponding "errors of observation". Similarly $y_1, y_2, y_3, \ldots\ldots$ is another series of true departures with $b_1, b_2, b_3, \ldots\ldots$ the corresponding errors. Thus the observed departures are $x_1' = x_1 + a_1$, $x'_2 = x_2 + a_2 \ldots\ldots$ and $y'_1 = y_1 + b_1$, $y'_2 = y_2 + b_2 \ldots\ldots$

3. I shall now proceed to deduce expressions corresponding to those of Chapman, but without making any assumption. We have if $x' = x + a$ and $y' = y + b$, $x'^2 = x^2 + 2ax + a^2$, $y'^2 = y^2 + 2by + b^2$, $x'y' = xy + ay + bx + ab$. Summing, dividing by the total number of records, and writing $s_x, s_y, s_{x'}, s_{y'}, s_a, s_3$ for the standard deviations of x, y, x', y', a and b respectively and $r_{xy}, r_{x'y'}, r_{ax}$, etc., for the coefficients of correlation between x and y, x' and y', a and x, etc., respectively, we get

3.

$$s_{x'}^2 = s_x^2 + 2\, s_a\, s_x\, r_{ax} + s_a^2$$

$$s_{y'}^2 = s_y^2 + 2\, s_b\, s_y\, r_{by} + s_b^2$$

$$s_{x'} . s_{y'} . r_{x'y'} = s_x . s_y\, r_{xy} + s_a . s_y\, r_{ay} + s_b . s_x . r_{bx} + s_a . s_b . r_{ab}$$

(1) Proc. Roy. Soc., A 98 (1920), pp. 235—248.

(2) M. O. No. 210-b., Geophysical Memoirs, No. 2.
M. O. No. 220-c., Geophysical Memoirs, No. 13, etc.

Hence,

$$r_{xy}=r_{x'y'}\cdot\frac{s_{x'}\,s_{y'}}{s_x\,s_y}-r_{ay}\cdot\frac{s_a}{s_x}-r_{bx}\cdot\frac{s_b}{s_y}-r_{ab}\cdot\frac{s_a\,s_b}{s_x\,s_y};$$

and, neglecting terms higher than the second,

$$r_{xy}=r_{x'y'}\left[1+r_{ax}\cdot\frac{s_a}{s_x}+r_{by}\frac{s_b}{s_y}+\tfrac{1}{2}\left\{(1-r_{ax}^2)\frac{s_a^2}{s_x^2}+(1-r_{by}^2)\frac{s_b^2}{s_y^2}+(2r_{ax}\,r_{by}\frac{s_a\,s_b}{s_x\,s_y}\right\}\right]$$

$$-r_{ay}\frac{s_a}{s_x}-r_{bx}\frac{s_b}{s_y}-r_{ab}\frac{s_a\,s_b}{s_x\,s_y}\quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

4. Chapman makes five different assumptions, (*i*) $r_{ab}=0$, (*ii*) $r_{ax}=0$, (*iii*) $r_{by}=0$, (*iv*) $r_{ay}=0$ and (*v*) $r_{bx}=0$. From a statistical point of view each of these assumptions requires justification.

I shall now investigate r_{ab} in greater detail, taking the particular ascent described in Computer's Handbook M. O. 223, Section 2, pp. 18—21, 26—28, and 40—42 as a concrete example.

5. The height H'_n, corresponding to the observed pressure and temperature P'_n and T'_n, is calculated by the successive application of Laplace's formula; the difference in height h'_{n-1} between H'_{n-1} and H'_n is given by

$$\log P'_{n-1}-\log P'_n=\frac{14{\cdot}837}{T_m}(1-{\cdot}00259\ \mathrm{Cos}\ 2\lambda)\ (1-H'/R)\ h'_{n-1}/1000 \quad . (A)$$

where P'_{n-1} = pressure at H'_{n-1}, T_m = mean temperature between H'_{n-1} and H'_n, γ=latitude and H'=mean height=$\frac{1}{2}(H'_n+H'_{n-1})$ and R=the terrestrial radius.

Putting $\quad T_m=(T'_n+T'_{n-1})/2$

and

$$14{\cdot}837\ (1-{\cdot}00259\ \mathrm{Cos}\ 2\lambda)\ (1-H'/R)=f(\lambda,H')$$

we have in kilometres

$$h'_{n-1}=(T'_n+T'_{n-1})\ (\log P'_{n-1}-\log P'_n)/2f(\lambda,H')\ . \quad . \quad . \quad . \quad (B)$$

This is the "observed" step, the corresponding "true" step being

$$h_{n-1}=(T_n+T_{n-1})\ (\log P_{n-1}-\log P_n)/2f(\lambda,H)$$

where T_n, T_{n-1}, P_n, P_{n-1} are "true" values.

6. Let $T'_n-T_n=dT_n$, $T'_{n-1}-T_{n-1}=dT_{n-1}$, $P'_n-P_n=dP_n$, $P'_{n-1}-P_{n-1}=dP_{n-1}$, etc., be "errors of observation" produced by (*i*) lag of the recording apparatus, (*ii*) errors in micrometer and stage readings of the trace, and (*iii*) errors in calculation.

Then

$$h_{n-1}=\left[(T'_n+T'_{n-1}-dT_n-dT_{n-1})\left\{\log P'_{n-1}\left(1-\frac{dP_{n-1}}{P_{n-1}}\right)-\log P'_n\left(1-\frac{dP_n}{P'_n}\right)\right\}\right]/2f(\lambda,H')$$

$$=\frac{(T'_n+T'_{n-1})}{2f(\lambda,H')}(\log P'_{n-1}-\log P'_n)\left\{\left(1-\frac{dT_n+dT_{n-1}}{T_n+T_{n-1}}\right)\left(1-\frac{\frac{dP_{n-1}}{P'_{n-1}}-\frac{dP_n}{P'_n}}{\log P'_{n-1}-\log P'_n}\right)\right\}\ .$$

Neglecting the very small difference between $f(\lambda, H')$ and $f(\lambda, H)$

$$h_{n-1} = h'_{n-1}\left\{1 - \frac{dT_n + dT_{n-1}}{T'_n + T'_{n-1}} - \frac{(dP_{n-1}/P'_{n-1}) - (dP_n/P'_n)}{\log P'_{n-1} - \log P'_n}\right\} \quad . \quad (2\cdot1)$$

$$= h'_{n-1}(1 - c'_{n-1}), \textit{ say}$$

where

$$c'_{n-1} = \frac{dT_n + dT_{n-1}}{T'_n + T'_{n-1}} + \frac{(dP_{n-1}/P'_{n-1}) - (dP_n/P'_n)}{\log P'_{n-1} - \log P'_n} \quad . \quad . \quad . \quad (2\cdot2)$$

The observational error in h_{n-1} is given by

$$dh_{n-1} = h'_{n-1} - h_{n-1} = + c'_{n-1} . h'_{n-1}$$

Now

$$H_n = H_{n-1} + h_{n-1} = H_{n-2} + h_{n-2} + h_{n-1} \quad . \quad . \quad . \quad .$$

$$= H_o + \overset{n-1}{\underset{o}{S}}(h_{n-1})$$

and

$$H'_n = H_o + \overset{n-1}{\underset{o}{S}}(h'_{n-1})$$

since H_o the height of the station is free from errors of observation. Therefore observational error in H_n is

$$dH_n = H'_n - H_n = \overset{n-1}{\underset{o}{S}}(h'_{n-1}) - \overset{n-1}{\underset{o}{S}}(h_{n-1}) = \overset{n-1}{\underset{o}{S}}(dh_{n-1})$$

$$= \overset{n-1}{\underset{o}{S}}(c'_{n-1} . h'_{n-1}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2\cdot3)$$

7. Let A′, B′ be two "observed" points on the trace corresponding to "true" points A, B. Let P'_n, T'_n, H'_n and P'_{n-1}, T'_{n-1}, H'_{n-1} be observed pressures, temperatures and heights of A′ and B′ respectively. Let a complete kilometer, say 9 km. be situated between A′ and B′. Then the pressure and temperature at 9 km. is found by linear interpolation between A′ and B′.

The observed values are

$$P'_9 = P'_n + \frac{H'_n - 9}{H'_n - H'_{n-1}} . (P'_{n-1} - P'_n)$$

$$T'_9 = T'_n + \frac{H'_n - 9}{H'_n - H'_{n-1}} . (T'_{n-1} - T'_n)$$

while corresponding "true" values are given by

$$P_9 = P_n + \frac{H_n - 9}{H_n - H_{n-1}} . (P_{n-1} - P_n)$$

and

$$T_9 = T_n + \frac{H_n - 9}{H_n - H_{n-1}} . (T_{n-1} - T_n).$$

Thus, adopting Chapman's notation, "errors of observation" in P_9 and T are given by

$$a_9 = T'_9 - T_9 = dT_n + \frac{H'_n-9}{H'_n-H'_{n-1}}\,(T'_{n-1}-T'_n) - \frac{H_n-9}{H_n-H_{n-1}}\,(T_{n-1}-T_n)$$

$$b_9 = P'_9 - P_9 = dP_n + \frac{H'_n-9}{H'_n-H'_{n-1}}\,(P'_{n-1}-P'_n) - \frac{H_n-9}{H_n-H_{n-1}}\,(P_{n-1}-P_n)$$

We get

$$a_9 = dT_n + \frac{(H'_n-9)(T'_{n-1}-T'_n)}{H'_n-H'_{n-1}}\left[1-\frac{(1-dH_n/H'_n-9)\,\{1-(dT_{n-1}-dT_n)/(T'_{n-1}-T'_n)\}}{1-(dH_n-dH_{n-1})/(H'_n-H'_{n-1})}\right]$$

$$= dT_n + \left(\frac{H'_n-9}{H'_n-H'_{n-1}}\right)\left(T'_{n-1}-T'_n\right)\left[\frac{dH_n}{H'_n-9}+\frac{dT_{n-1}-dT_n}{T'_{n-1}-T'_n}-\frac{dH_n-dH_{n-1}}{H'_n-H'_{n-1}}\right]$$

neglecting terms of the second order.

But $\dfrac{dH_n-dH_{n-1}}{H'_n-H'_{n-1}} = \dfrac{(H'_n-H_n)-(H'_{n-1}-H_{n-1})}{H'_n-H'_{n-1}} = \dfrac{h'_{n-1}-h_{n-1}}{h'_{n-1}} = c'_{n-1}$

Writing $\dfrac{H'_n-9}{H'_n-H'_{n-1}} = f'_n$, we get

$$a_9 = dT_n + f'_n\,.\,(T'_{n-1}-T'_n)\left\{\frac{dH_n}{H'_n-9}+\frac{dT_{n-1}-dT_n}{T'_{n-1}-T'_n}-c'_{n-1}\right\} \quad\text{.........(3·1)}$$

$$b_9 = dP_n + f'_n\,(P'_{n-1}-P'_n)\left\{\frac{dH_n}{H'_n-9}+\frac{dT_{n-1}-dT_n}{T'_{n-1}-T'_n}-c'_{n-1}\right\} \quad\text{.........(3·2)}$$

8. It will be now necessary to make definite assumptions about the quantities dP_n, dP_{n-1}, dT_n, dT_{n-1}, etc.

Case I.—We assume that $dT_n = dT_{n-1} \ldots = dT$ and $dP_n = dP_{n-1} = \ldots dP$, *i.e.*, assume that owing to some mechanical defect the errors are constant throughout the whole of the trace.

Then

$$a_9 = dT + f'_n\,(T'_{n-1}-T'_n)\left(\frac{dH_n}{H'_n-9}-c'_{n-1}\right)$$

$$b_9 = dP + f'_n\,(P'_{n-1}-P'_n)\left(\frac{dH_n}{H'_n-9}-c'_{n-1}\right)$$

where

$$c'_{n-1} = \frac{2\,dT}{T'_n + T'_{n-1}} + \frac{dP\,\{(1/P'_{n-1}) - (1/P'_n)\}}{\log P'_{n-1} - \log P'_n}.$$

and $dH_n = \overset{n-1}{\underset{0}{S}}\,(c'_{n-1}.h'_{n-1})$, or from (*B*), p. 12,

$$= \frac{1}{f(\gamma, H)}\left[dT\,(\log P'_0 - \log P'_n) + dP.\overset{n-1}{\underset{0}{S}}\left(\frac{T'_n + T'_{n-1}}{2}\right)\left(\frac{1}{P'_{n-1}} - \frac{1}{P'_n}\right)\right]$$

We can easily find the value of this expression for the ascent described in the Computer's Handbook M. O. 223 § 2, p. 41, etc.

$$H'_n = 9{\cdot}327,\quad P'_n = 289,\quad T'_n = 220,\quad f(\lambda, H) = 14{\cdot}82$$

$$H'_{n-1} = 8{\cdot}764,\quad P'_{n-1} = 315,\quad T'_{n-1} = 225,\qquad f'_n = {\cdot}5808$$

Also from table given on p. 41 we compute

$$\overset{n-1}{\underset{0}{S}}\left(\frac{T'_n + T'_{n-1}}{2}\right)\left(\frac{1}{P'_{n-1}} - \frac{1}{P'_n}\right) = 0{\cdot}6113$$

Thus

$$\frac{dH_n}{H'_n - 9} = +\,0{\cdot}1105\;dT - 0{\cdot}1262\;dP$$

$$c'_{n-1} = +\,0{\cdot}004494\;dT - {\cdot}007637\;dP$$

Hence

$$a_9 = +\,1{\cdot}3078\;dT - 0{\cdot}3483\;dP$$

and

$$b_9 = +\,1{\cdot}6007\;dT - 0{\cdot}7904\;dP$$

Multiplying and summing

$$S(a_9.b_9) = +\,2{\cdot}09\;S(dT^2) + 0{\cdot}28\;S(dP^2)\ldots\ldots\ldots\ldots\ldots\ (4)$$

even when $S(dP.dT) = 0$. Thus even when errors dP and dT are uncorrelated the interpolation errors a_9 and b_9 will not be uncorrelated.

Again

$$S(x.a) = +\,1{\cdot}3078\;S(x.dT) - 0{\cdot}3483\;S(x.dP)$$

Thus r_{ax} will not in general be zero unless x and dT, and x and dP are both uncorrelated. Hence if the observational error produced by the lag of the instrument depend at all in the degree of deviation from normal, then r_{ax} (and r_{by}, r_{ay}, r_{bx}) will not be zero.

9. We now assume that $\dfrac{dT_n}{T'_n} = \dfrac{dT_{n-1}}{T'_{n-1}} = \ldots\ldots\ t'$ and $\dfrac{dP_n}{P'_n} = \dfrac{dP_{n-1}}{P'_{n-1}} = \ldots\ p'$

That is, the errors are proportional to the pressures and temperatures.

Then

$$a_9 = t'.T'_n + f'_n\;(T'_{n-1} - T'_n)\;\left\{\frac{dH_n}{H'_n - 9} + t' - c'_{n-1}\right\}$$

$$b_9 = p'.P'_n + f'_n.(P'_{n-1} - P'_n)\;\left\{\frac{dH_n}{H'_n - 9} + p' - c'_{n-1}\right\}$$

But in this case $c'_{n-1} = t'$, and

$$dH_n = \overset{n-1}{\underset{o}{S}} (c'_{n-1}.h_{n-1}) = t'.\overset{n-1}{\underset{o}{S}} (h_{n-1}) = t'(H'_n - H_o)$$

Thus we find, since $H_o = \cdot 060$

$$a_9 = +\ 302\cdot 3\ t' \text{ and } b_9 = +\ 304\cdot 1\ p' + 412\cdot 83\ t'$$

Therefore, $S(a_9 \cdot b_9) = +\ 124798\cdot 5\ S(t^2)$ ....(5)
even if p' and t' are uncorrelated, that is r_{ab} is not zero.

10. Assuming $dT_n = dT_{n-1} \cdots\cdots = dT$, but $\dfrac{dP}{P'_n} = \dfrac{dP_{n-1}}{P'_{n-1}} \cdots\cdots\cdots = p'$.

$$c'_{n-1} = 2\ dT/(T'_{n-1} + T'_n)$$

$$dH_n = \overset{n-1}{\underset{o}{S}} \left(\frac{2\ dT}{T'_{n-1}+T'_n}\right) \left\{ \frac{T'_n + T'_{n-1}}{2\ f(\lambda, H)} \left(\log P'_{n-1} - \log P'_n\right) \right\}$$

$$= S\ dT. (\log P'_{n-1} - \log P'_n)\ / f(\lambda,\ H) = +0\cdot 03614\ dT$$

Therefore $a_9 = +1\cdot 3078\ dT$ and $b_9 = +304\cdot 1\ p' + 1\cdot 6\ dT$

Thus $S(a_9 . b_9) = +2\cdot 09\ S(dT^2)$ ... ... (6)

even if p' and dT are uncorrelated.

11. Assuming $\dfrac{dT'_n}{T'_n} = \dfrac{dT_{n-1}}{T'_{n-1}} = \cdots\cdots\cdots t'$ while $dP_n = dP_{n-1} \cdots\cdots\cdots = dP$

we get

$$c'_{n-1} = t' + \frac{dP \left\{ \dfrac{1}{P'_{n-1}} - \dfrac{1}{P'_n} \right\}}{\log P'_{n-1} - \log P'_n}$$

$$dH_n = S\left(t'.\ h'_{n-1}\right) + \frac{dP}{2f(\lambda, H)} . S\left(T'_n + T'_{n-1}\right)\left(\frac{1}{P'_{n-1}} - \frac{1}{P'_n}\right)$$

$$= +9\cdot 267\ t' + 0\cdot 04125\ dP$$

Hence $a_9 = +302\cdot 3\ t' - \cdot 04125\ dP$ and $b_9 = +412\cdot 83\ t' - 0\cdot 7904\ dP$

Therefore $S(a_9 . b_9) = +124798\cdot 5\ S(t^2) + 0\cdot 28\ S(dP^2)$ ... ... (7)

when t' and dP are uncorrelated.

12. I have investigated four different cases, assuming (*i*) dP, dT constant, (*ii*) dP/P and dT/T constant, (*iii*) dP/P and dT constant, (*iv*) dT/T and dP constant respectively. In each case we found that r_{ab} is not zero even when observational

errors are uncorrelated, and it seems difficult to avoid the expectation that in as far as errors are due to mechanical causes it will very rarely, if ever, happen that r_{ab} vanishes.

13. From the form of the equations (4) to (7) high values of r_{ab} do not seem improbable. In equation (1) assuming r_{ay}, r_{bx}, r_{ax} and r_{by} to be each zero, we get

$$r_{xy} = r_{x'y'} \left\{ 1 + \tfrac{1}{2}\left(\frac{s_a^2}{s_x^2} + \frac{s_b^2}{s_y^2} \right) \right\} - \frac{s_a\, s_b}{s_x\, s_y} \cdot r_{ab} \qquad \ldots \qquad (1{\cdot}1)$$

$$= (1 + e_1)\, r_{x'y'} - e_2\, r_{ab}$$

$$= r_{x'y'} + (e_1\, r_{x'y'} - e_2\, r_{ab})$$

where $e_1 \equiv \tfrac{1}{2}\left(\frac{s_a^2}{s_x^2} + \frac{s_b^2}{s_y^2} \right)$ and $e_2 = \frac{s_a\, s_b}{s_x\, s_y}$

Hence

$$r_{xy} \lesseqgtr r_{x'y'} \text{ according as } r_{ab} \gtreqless \frac{e_1}{e_2} \cdot r_{x'y'}.$$

14. In the following table I have taken the winter values of s_a/s_x and s_b/s_y quoted by Chapman on p. 238 of the paper already cited. I get the average value of $e_1/e_2 = 1{\cdot}012$

km.	s_a/s_x	s_b/s_y	e_1	e_2	e_1/e_2
2	·334	·416	·1423	·1389	1·024
4	·284	·326	·0935	·0926	1·010
6	·260	·280	·0730	·0728	1·003
8	·302	·238	·0739	·0719	1·028
10	·344	·220	·0833	·0760	1·096

Thus $r_{xy} \lesseqgtr r_{x'y'}$ according as $r_{ab} \gtreqless 1{\cdot}012\, r_{x'y'}$

Hence we conclude that if r_{ab} is of the same order as $r_{x'y'}$ the correction for "observational errors" will be negligible, while if r_{ab} is substantially greater than r_{xy} the actual value of the "true" correlation may even be less than the observed value $r_{x'y'}$.

15. Douglas (1) has recently published values of the correlation between pressure and temperature at 10,000 feet, which are considerably lower than Dines's values. Douglas gets ·65 for a total of 550 observations a number much larger than that used by Dines in any one of his sets. But the most important point about

(1) C. K. M. Douglas. Quar. Jour. R. Met. Soc., Vol. XLVII, 1921, pp. 23—26.

Douglas's work is the fact that his observations "refer to actual heights above mean sea-level, and not to aneroid heights" (1). I presume that Douglas's work is thus free from the "interpolation correlation" arising from the use of Laplace's formula for the calculation of the heights. In other words if dP, dT are uncorrelated then a, b the "interpolated" errors will also remain uncorrelated.

The low values of $r_{x'y'}$ obtained by Douglas certainly suggest the possibility that in Dines's case the correlation r_{ab} due to "interpolation" may be so high that the "observed" correlations $r_{x'y'}$ are actually greater than the true coefficients r_{xy}.

16. It will be remembered that in the foregoing discussion we had assumed that r_{ax}, r_{by}, r_{ay}, and r_{bx} are all zero. It may not be out of place to say a few words with regard to this assumption.

Karl Pearson (2) has definitely established the reality of a "genuine correlation in the judgments of independent observers". (3) He says:—

> "astronomers have already found that the brightness of a star influences the personal equation. This in the language of the present writer produces a correlation of judgments ...... Hence if a number of observations were made on stars of varying magnitude, the judgment being a function of the magnitude, we should have a series of correlated errors ...... When every effort is made to eliminate large causes, ... there still remains a multitude of small causes which produce correlation. It might be possible in an *ideal* series still further to eliminate some of these, but in practical observation we have to take a given phenomenon as it is, and we cannot possibly abstract from it the whole of its characteristic atmosphere".

Then again :—

> "the errors of judgment of apparently independent observers are not as a rule independent. The immediate atmosphere of each single observation or of each short series of observations affects in a differential manner the factors of the personality," (4).

Pearson also noticed that the correlation even extended to observations at different times by different observers of the same physical quantity. (5)

17. The present case is not exactly parallel. But when it is remembered that the same observer is measuring both x, a, y and b at the same time and from the same trace, it would seem that conditions are still more favourable for producing correlations between the different quantities measured. The correlations between x, a, y, and b due to "personal equation", what Pearson calls "atmosphere", may not be large, but even small values would sensibly affect Chapman's corrections.

(1) C. K. M. Douglas, Quar. Jour. R. Met. Soc., Vol. XLVII, 1921, p. 25.

(2) Errors of Judgment, etc., Phil. Trans, Vol. 198 A (1902).

(3) The same paper, page 292.

(4) The same paper, page 270.

(5) The same paper, page 290.

Speaking generally it seems likely that x and a or y and b will be more highly correlated than x and b or y and a. Let us assume that r_{ay} and r_{bx} are each zero while r_{ax} and r_{by} are small but negative. A negative correlation merely implies proportionally smaller errors for large fluctuations and from a purely physical point of view is not impossible. Assuming $r_{ax}=r_{by}=-\cdot 2$, we have from equation (*1*)

$$r_{xy}=r_{x'y'}\left[1-\cdot 2\left(\frac{s_a}{s_x}+\frac{s_b}{s_y}\right)+\tfrac{1}{2}\left\{\cdot 96\left(\frac{s_a^2}{s_x^2}+\frac{s_b^2}{s_y^2}\right)+\cdot 08\,\frac{s_a s_b}{s_x s_y}\right\}\right]-\frac{s_a s_b}{s_x s_y}\,r_{ab}\quad \ldots \quad (1\cdot 2)$$

Taking numerical value of s_a/s_x and s_b/s_y already considered, the average value of

$\left(\frac{s_a}{s_x}+\frac{s_b}{s_y}\right)=\cdot 62$, of $\left(\frac{s_a^2}{s_x^2}+\frac{s_b^2}{s_y^2}\right)=\cdot 186$, of $\frac{s_a s_b}{s_x s_y}=\cdot 141$. Thus $r_{xy}=r_{x'y'}(1-\cdot 124+\cdot 098)-\cdot 141$

Thus $r_{xy}=\cdot 974\, r_{x'y'}-\cdot 141\, r_{ab}$. In this case the true value r_{xy} must be distinctly less than the observed value $r_{x'y'}$ even if r_{ab} is less than $r_{x'y'}$.

Thus even a very small "personal" or "reading" correlation between x and a or y and b will substantially alter the correction for errors of observation.

18. The above discussion appears to show that Chapman's corrections cannot be accepted as real without further investigation and arguments based on them must be regarded as doubtful.

19. In conclusion I wish to acknowledge my indebtedness to Dr. Gilbert T. Walker, F.R.S., Director-General of Observatories, for suggesting the problem, for giving me permission to work in his department and for the help and encouragement he has given me at all times.

SGPI-21-IV-5(6)-11-1-23—400.