ART. XVI.—*Further Remarks on a method of Reducing Observations of Temperature;* by Professor J. D. EVERETT, of Kings College, Windsor, Nova Scotia.

IN an article in the January number of this Journal, I recommended for general use a method of comparing climates, as regards temperature, by reference to the curve whose equation is

$$y=A_0+A_1 \sin(x+E_1)$$

Professor Loomis, in his "*Remarks*" appended to my article, having impugned the accuracy of this mode of procedure, on the ground that the curve of temperature at a place does not possess the symmetry which characterizes the curve expressed by the above equation, I propose, in the present paper, to furnish a thorough investigation of the degree of accuracy which is attained.

We must first prove the following proposition: Any m numbers, (m being either $2n$ or $2n+1$,) can be *exactly* expressed by an equation of the form

$$y=A_0+A_1 \sin(x+E_1)+A_2 \sin(2x+E_2)+ \dots +A_n \sin(nx+E_n),$$

in such a sense that, by giving x the successive values,

$$0, \frac{1}{m}\times 360^\circ, \frac{2}{m}\times 360^\circ \dots \frac{m-1}{m}\times 360^\circ,$$

the resulting values of y will be the given numbers in order. We shall hereafter denote the terms on the second side of the above equation, by the symbols t_0, t_1, t_2, &c. To prove the proposition, let the series be transformed, (see p. 25 of my former article,) into

$$y=A_0+P_1 \cos x+Q_1 \sin x+P_2 \cos 2x+Q_2 \sin 2x+ \dots +P_n \cos nx+Q_n \sin nx.$$

We can then form m equations, by writing the given numbers in order for y, and giving x the corresponding values above indicated. If $m=2n+1$, we shall have $2n+1$ equations, to determine the $2n+1$ constants A_0, P_1, Q_1, &c. If $m=2n$, the last term, $Q_n \sin nx$, is to be omitted; and we shall have $2n$ equations to determine $2n$ constants. Hence, by the theory of equations, there will always be one and only one solution.

The rule for obtaining the solution is extremely simple:—

To find the value of any one of the constants, multiply each equation by its coefficient of that constant, and add the m equations thus obtained.[1] It will be found that all the terms on the

[1] It is worthy of remark that this rule is identical with that laid down, in the theory of probabilities, for determining the most probable values of any number of unknown quantities from a greater number of simple equations, when the latter are all of equal weight. This general rule is thus given in "*Airy on Errors of Observations,*" p. 80: "Multiply every equation by its coefficient of one unknown quantity, and take the sum for a new equation; the same for the second unknown quantity; and so on for every unknown quantity; and thus a number of equations will be found equal to the number of unknown quantities."

second side destroy one another, except those which contain the constant we are seeking, and the sum of these will be $\frac{m}{2}$ times the constant, except when the constant is P_n in the case where $m=2n$, for which the sum will be mP_n. When there are 12 given numbers, this process resolves itself into that described in my former article,[2] (*this Journal*, [2], xxxv, 17), supplemented by the following formulæ:

$$6P_5=k_0-s_2(k_1-k_5)+s_1(k_2-k_4)$$
$$6Q_5=\quad s_1(k_1-k_5)-s_2(k_2-k_4)+k_3$$
$$12P_6=K_0+K_2+K_4-K_1-K_3-K_5$$

The monthly means at Greenwich, derived from the table in Professor Loomis' "*Remarks,*" reckoning the 31st of January and 1st of March as part of February, are

36·6 38·6 41·4 46·2 52·9 59·0 61·8 61·0 56·5 49·9 43·2 39·3

These are exactly expressed by the formula,

$$y=48.87+12{\cdot}44\sin(x+262^\circ\ 31')+{\cdot}84\sin(2x+57^\circ\ 44')+{\cdot}18\sin(3x+338^\circ\ 12')$$
$$+{\cdot}25\sin(4x+253^\circ\ 26')+{\cdot}20\sin(5x+252^\circ\ 29')+{\cdot}13\sin(6x+270^\circ),$$

where x is 0° for January, 30° for February, 60° for March, and so on.

Here t_0 is the constant 48·87.

The values of t_1, for the 12 months in order, are

−12·33 −11·49 −7·57 −1·62 +4·76 +9·87 +12·33 +11·49 +7·57 +1·62 −4·76 −9·87.

The values of t_2 are

+·71 +·74 +·03 −·71 −·74 −·03 +·71 +74 +·03 −·71 −74 −·03

These values, it will be observed, repeat themselves after the first six; and the sum of any consecutive six is 0.

The values of t_3 are

−·07 +·17 +·07 −·17 −·07 +·17 +·07 − ·17 −·07 +·17 +·07 −·17

which repeat themselves after the first four, or go through their cycle in a third of a year. Also the sum of any consecutive four is 0.

The values of t_4 are

−·24 +·06 +·18 −·24 +·06 +·18 −·24 +·06 +·18 −·24 +·06 +·18

which go through their cycle 4 times in the year. Also the sum of any consecutive three is 0.

The values of t_5 are

−·19 +·13 −·04 −·06 +·15 −·19 +·19 −·13 +·04 +·06 −·15 +·19

[2] The following errata in the formulæ there given require correction:—
p. 26, lines 9, 10, 11, for K_0+K_3 read K_0-K_3
" K_1+K_4 " K_1-K_4
" K_2+K_5 " K_2-K_5

which do not so plainly show the recurrence of cycles, because 5 is not a measure of 12. It is easy however to trace 5 maxima and 5 minima.

The values of t_6 are

$-\cdot13 + \cdot13 - \cdot13 + \cdot13 - \cdot13 + \cdot13 - \cdot13 + \cdot13 - \cdot13 + \cdot13 - 13 + \cdot13$

which go through their cycle 6 times in the year. Also the sum of any consecutive two is 0.

By addition, we find the values of $t_0+t_1+t_2+t_3+t_4+t_5+t_6$ to be

36·62	38·61	41·41	46·20	52·90	59·00	61·80	60·99	56·49	49·90	43·22	39·30

or, to one place of decimals,

36·6	38·6	41·4	46·2	52·9	59·0	61·8	61·0	56·5	49·9	43·2	39·3

which agree exactly with the monthly means.

The error committed by stopping at any term, is, of course, equal to the sum of the terms omitted. These sums are given in the following table; the first line containing the sum of all the terms after t_1, the second line the sum of all after t_2, and so on,—

+·08	+1·23	+·11	−1·05	−·73	+·26	+·60	+·63	+·05	−·59	−·87	+·30
−·63	+ ·49	+·08	− ·34	+·01	+·29	−·11	−·11	+·02	+·12	−·13	+·33
−·56	+ ·32	+·01	− ·17	+·08	+·12	−·18	+·06	+·09	−·05	−·20	+·50
−·32	+ ·26	−·17	+ ·07	+·02	−·06	+·06	+·00	−·09	+·19	−·26	+·32
−·13	+ ·13	−·13	+ ·13	−·13	+·13	−·13	+·13	−·13	+·13	−·13	+·13

The terms t_2, t_4, t_6 produce no effect upon the mean temperature of a half-year, for the sum of any 6 consecutive values of these terms is 0; hence, in deriving the mean temperature of a half-year from t_0+t_1, the error committed depends only on t_3 and t_5. The sums of t_3 and t_5, for each month, are

$-\cdot26 + \cdot30 + \cdot03 - \cdot23 + \cdot08 - \cdot02 + \cdot26 - \cdot30 - \cdot03 + \cdot23 - \cdot08 + \cdot02$

and the means of every consecutive six of these, commencing with October—March, are

$+\cdot04 - \cdot04 - \cdot01 - \cdot02 + \cdot07 - \cdot03 - \cdot04 + \cdot04 + \cdot01 + \cdot02 - \cdot07 + \cdot03$

which are the corrections required in determining the mean temperature of 6 consecutive calendar months from t_0+t_1. In fact, the means of t_0+t_1 for every 6 months, in the above order, are

41·47	40·93	42·52	45·81	49·92	53·75	56·27	56·81	55·22	51·93	47·82	43·99

while the means of the actual monthly temperatures are

41·50	40·88	42·50	45·78	49·98	53·72	56·23	56·85	55·23	51·95	47·75	44·02

showing the same corrections as above, allowing differences of 1 in the last figure for neglected places of decimals. Hence we see that the mean temperature of any 6 consecutive months at Greenwich can be calculated from t_0 and t_1 (or from their constants A_0 A_1 E_1), subject to errors not exceeding ·07 of a degree.

It may be shown in general, by a theorem proved on p. 30 of my former article, that, if the expression for monthly means be

$$y = t_0 + t_1 + t_2 + t_3 + t_4 + t_5 + t_6,$$

the expression for half-yearly means will be

$$Y = t_0 + {\cdot}644 t_1 - {\cdot}236 t_3 + {\cdot}173 t_5 ;$$

that is,

$$Y = A_0 + {\cdot}644\, A_1 \sin(x+E_1) - {\cdot}236 A_3 \sin(3x+E_3) + {\cdot}173 A_5 \sin(5x+E_5)$$

and since $\cdot 236\, A_3$ and $\cdot 173\, A_5$ are practically insignificant compared with $\cdot 644\, A_1$, the two last terms may be neglected.

For Greenwich, we have

$$\begin{aligned} {\cdot}644 A_1 &= {\cdot}644 \times 12{\cdot}44 = 8{\cdot}01, \\ {\cdot}236 A_3 &= {\cdot}236 \times \;\; {\cdot}18 = \;\; {\cdot}04, \\ {\cdot}173 A_5 &\doteq {\cdot}173 \times \;\; {\cdot}20 = \;\; {\cdot}03. \end{aligned}$$

Hence the sum of the two last terms in the expression for Y cannot exceed $\cdot 04 + \cdot 03 = \cdot 07$, which is only $\frac{1}{114}$ of $\cdot 644\, A_1$; and the greatest error produced by the omission of these terms in comparing two halves of the year cannot be more than about $\frac{1}{114}$ of the difference between the greatest and least values of Y. It is to be observed that the half-yearly means denoted by Y are not limited to means of 6 calendar months, but may belong to any 182½ consecutive days.

It thus appears that the values of A_0, A_1, E_1 are sufficient to determine, with competent accuracy, the mean temperature of any half-year, and hence to determine the range as measured by the difference between the warmest and coldest halves of the year, a system of measurement which I think is as fair as any that can be devised.

The determination of the precise dates of maximum and minimum is always difficult, whatever method be pursued, owing to the slow rate at which temperature changes when near its maximum or minimum. Even the table of daily temperatures at Greenwich, though derived from an average of 43 years, leaves both these dates doubtful, as a glance will show. It is always much easier to determine the dates at which the temperature is equal to the mean of the year, (or at which the curve of temperature intersects the line of mean annual temperature,) because at and near these dates the temperature changes with its greatest rapidity. This remark applies to half-yearly means, as well as to daily temperatures; hence we can determine more precisely "the dates which divide the year into halves whose temperatures are each equal to the mean of the year," than the dates which are the centres respectively of the warmest and coldest halves of the year. For the former determination depends upon the dates at which the value of Y is equal to the mean annual temperature, while the latter depends upon the dates at which Y is a maximum or minimum. And the above phrase in inverted commas may be advantageously substituted for "centres of warm and cold halves of the year" in the definitions of E_1 in my former article.

We will now proceed to test our determinations of range and date by the Greenwich table.

The range, as measured by the difference between the warmest and coldest halves of the year, is by our theory equal to $A_1 \times 1{\cdot}2879 = 12{\cdot}44 \times 1{\cdot}2879 = 16{\cdot}05$. This is in error by about ·07, the warmest 182 days, April 22—Oct. 30, having a mean temperature of 57·00, and the coldest 183 days, Oct. 31—April 21, a mean temperature of 40·88, showing a difference of 16·12.

As regards date, the value of E_1 is 262° 31′ or 82° 31′, according as we make A_1 positive or negative. Subtracting 15° from the latter, to reduce to the beginning of the year, we have for remainder 67° 31′, the complement of which is 22° 29′, equivalent to 22·8 days; and half a year or 182·5 days, added to this, gives 205·3 days. The 22d and 205th days of the year are January 22 and July 24; hence, by our theory, January 22·8 and July 24·3 should divide the year into halves whose mean temperatures are equal to each other and to the mean annual temperature.

From the Greenwich table we derive the following half-yearly means:—

Jan. 23	—July 23, both inclusive,	182 days, mean temp.	49·0
23	24	183	49·2
22	22	182	48·8
22	23	183	48·9

the mean annual temperature being 48·9. Hence our determination of date is only 1 day in error.

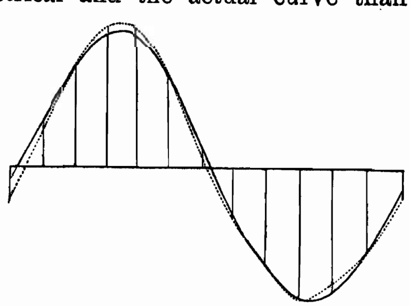
The centres of the warmest and coldest halves, as already stated, cannot be determined with so much precision. In the present case, the coldest 183 days are October 21—April 21, the centre of which is January 20, whereas our theory makes it January 22·8.

We have a good approximation, in the present case, to the dates which are midway between those whose temperatures are the same as the mean of the year. The latter, (which we may call the Vernal and Autumnal intersections,) are April 29 and October 21; and the dates midway between these are January 23·5 and July 25, which agree closely with January 22·8 and July 24·3. I believe this agreement will be generally found to exist, because the term t_1 is changing its value with nearly maximum rapidity at the dates of intersection, and thus overpowers the other terms, especially as t_2, the most considerable of them, cannot hasten or retard one of the intersections to any considerable extent, without producing an opposite and nearly equal effect upon the other.

The influence of the term t_2, in modifying the symmetrical curve $y=t_0+t_1$ for Greenwich, may be thus described:—It retards the vernal intersection, and the maximum; hastens the autumnal intersection, and the minimum; intensifies the maximum, and moderates the minimum. Hence the temperature rises higher above the mean than it falls below; but the number of days below the mean is greater than the number above. These features of climate will probably be found to prevail generally in extra-tropical latitudes; and they constitute the most marked departures of the curve of temperature from symmetry, the terms after t_2 being comparatively insignificant. In so far as they are common to all places, they are unimportant in the comparison of climates, which was the object proposed in my former paper; and, at all events, the first elements that claim attention are those in which the actual curve agrees with the symmetrical curve, that is to say, the elèments of mean temperature, range, and general earliness, as represented by the three constants A_0 A_1 and E_1.

Of course, three numbers cannot express every feature of the curve of temperature at a place, for it is in the nature of things impossible that three numbers should express more than three independent elements. If the constants A_0 A_1 and E_1 are correct as measures of the three leading elements, it is all that we can expect of them; and if a closer approximation to the curve of temperature is desired, it can readily be obtained by computing the constants in one or more of the succeeding terms.

The subjoined diagram will give a clearer impression of the relation between the symmetrical and the actual curve than language can convey. The continuous line is the symmetrical curve for Greenwich; and the dotted line is a curve, drawn through points whose ordinates are the mean temperatures of the 12 months, beginning and ending with April, these points being situated on the vertical lines in the diagram.



The horizontal line represents the mean annual temperature. Of the areas intercepted between the two curves, those which lie above either of the curves are together equal to those which lie below the same.