Relation between growth and yield of the Sugarcane Crop at Poona

R. P. SARKER, V. K. RAGHAVENDRA and P. S. NAYAR

Meteorological Office, Poona

(*Received 17 March 1962*)

1. Introduction

The growth phenomena like germination, tillering, elongation, ripening etc and the ultimate yield of a crop are influenced by the weather elements, which the crop experiences during its life-cycle. Accordingly, it may be possible to find out a regression formula between the yield and the weather elements (experienced by the crop during its life-span) and to predict the yield therefrom. However, "one obvious limitation to a prediction formula based on meteorological variates is that the weather vagaries right up to the time of harvest would probably have to be included. An alternative is to use the plant itself as an integrator of weather effects and to base a prediction formula for yield on easily recognisable and measurable states of the crop during its growth" (Keen 1940). It may, therefore, be possible to forecast yield by working out a direct regression between the yield and the crop characteristics such as germination, tillering, maximum height etc. For such studies, however, records extending over a long series of years will be essential.

In the present study an attempt (based on 14 years' data) has been made to estimate the yield of sugarcane at Poona from the periodical observations of the height of the crop taken from the beginning and extending up to the time the maximum height is attained. The crop at Poona is harvested by the end of January, whereas it attains its maximum height by the end of November. The present study thus enables the prediction of the yield two months in advance of the harvest.

2. Material and method

Systematic observations on the developmental characteristics are being recorded on sugarcane under the Co-ordinated Crop-Weather Scheme. These observations are taken on 72 clumps spread over the entire field, in six growth observation plots, selected at random according to the sampling techniques laid down in the All India Co-ordinated Crop-Weather Scheme. Average weekly values of progressive height measurements, according to this technique, are available for all the years, beginning from 1946-47. The periodical quantitative measurements on the growing crop have to be fitted by a curve. Ramabhadran (1950) found that the best curve to the growth data is the skew-logistic curve which is based on the assumption that the rate of growth at any time depends on (*i*) the growth already attained (accelerating factor), (*ii*) growth yet to be attained (inhibiting factor) and (*iii*) a function of the age of the crop. The curve is given by $dy/dx = F(x)y\ (K\text{-}y)$ where y is the progressive growth at the time x and K, the maximum height to be attained. On integration, the skew-logistic curve takes the form

$$y = d + \frac{K}{1 + e^{a_0 + a_1x + a_2x^2 + a_3x^3}} \qquad (1)$$

keeping up to 3rd power of x. Since this curve fits the observations closely it may be said the set of six constants K, d, a_0 a_1, a_2, a_3 represent the entire growth in height. These

TABLE 1

Growth constants for height values of variety POJ. 2878 Sugarcane at Poona

Year	$K + d$	a_0	a_1	a_2	a_3
1946-47	452·9	5·1286	—0·5306	0·0212	—0·00040
1947-48	400·7	5·2697	—0·4103	0·0149	—0·00027
1948-49	354·0	5·8408	—0·4654	0·0161	—0·00027
1949-50	401·9	5·0265	—0·3345	0·0063	—0·00008
1950-51	353·0	5·1960	—0·3539	0·0103	—0·00022
1951-52	334·4	5·0857	—0·4066	0·0122	—0·00021
1952-53	332·5	5·1898	—0·3087	0·0059	—0·00009
1953-54	319·2	4·2301	—0·2783	0·0070	—0·00015
1954-55	306·4	5·1650	—0·3184	0·0057	—0·00081
1955-56	359·3	5·1126	—0·5654	0·0255	—0·00042
1956-57	365·2	5·9647	—0·4132	0·0114	—0·00021
1957-58	315·9	4·6260	—0·2749	0·0048	—0·00010
1958-59	356·4	5·0756	—0·3207	0·0049	—0·00007
1959-60	316·1	5·2986	—0·4379	0·0167	—0·00033

constants may be termed "Growth Constants". In order, thus, to predict the yield from the periodical observations of height values a regression formula between yield and these growth constants can be utilised. In the present study this skew-logistic curve has been fitted to the height data of one of the varieties of sugarcane, *viz.*, POJ. 2878 grown under the Co-ordinated Crop-Weather Scheme, at Poona, and the values of the growth constants have been found out for individual years and are given in Table 1.

With the help of these constants a regression formula to forecast yield has been worked out.

The method of fitting the skew-logistic curve (1) to the progressive height values and of finding the "growth constants" is given below for a particular year 1960-61 as an example.

3. Fitting the Skew-logistic curve to the 1960-61 data

I. *First approximation*

The skew logistic curve

$$y = d + \frac{K}{1+e^{a_0+a_1x+a_2x^2+a_3x^3}} \qquad (2)$$

has the two asymptotes $y=d$ and $y=K+d$.

The weekly progressive height values y are first plotted against time x and a smooth free-hand curve is drawn. The lower asymptote $y=d$ is then fixed from this curve.

Shifting the origin to $(0, d)$ the equation takes the more convenient form

$$y' = y - d = \frac{K}{1+e^{a_0+a_1x+a_2x^2+a_3x^3}}$$

$$\text{or, } a_0 + a_1x + a_2x^2 + a_3x^3 = \log_e \frac{K-y'}{y'} \qquad (3)$$

The constants K, a_0, a_1, a_2, a_3 are found (Davies *vide* Ref.) as follows—

Five equidistant points $(0, y_0)$, (x_1, y_1), $(2x_1, y_2)$, $(3x_1, y_3)$, and $(4x_1, y_4)$ representative of the different portions of the curve, are chosen from the free-hand curve. Substituting these points in equation (3) and eliminating the constants a_0, a_1, a_2, a_3 by algebraic simplification, the constant K is found from the equation

$$\frac{(K-y_0)\,(K-y_2)^6\,(K-y_4)}{(K-y_1)^4\,(K-y_3)^4} = \frac{y_0\,y_2^6\,y_4}{y_1^4\,y_3^4} \qquad (4)$$

$$(K \neq 0)$$

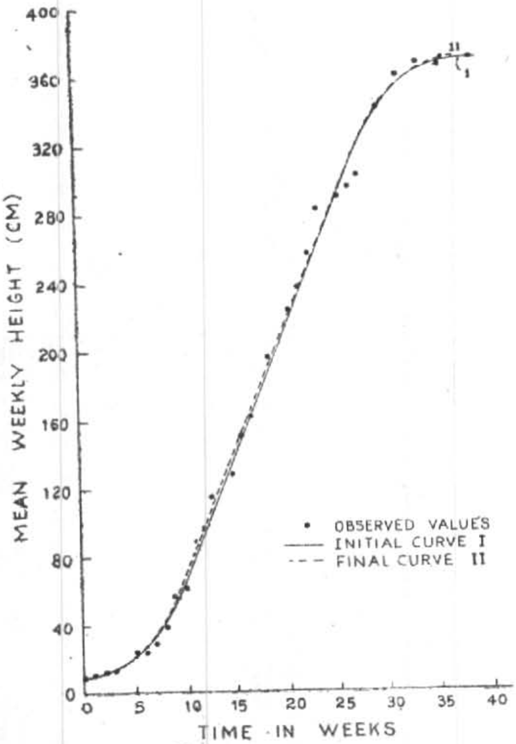

Fig. 1. Poona, Sugarcane POJ. 2878 (1960-61)

The other constants are found from the equations—

$$\left.\begin{aligned} a_0 &= \log_e \frac{K-y_0}{y_0} \\ a_1 &= \frac{18\beta_1-9\beta_2+2\beta_3}{6x_1} \\ a_2 &= \frac{4\beta_2-5\beta_1-\beta_3}{2x_1^2} \\ a_3 &= \frac{\beta_3+3\beta_1-3\beta_2}{6x_1^3} \end{aligned}\right\} \quad (5)$$

where the β values are obtained from the equations :

$$\left.\begin{aligned} \beta_1 &= \log_e \frac{y_0\,(K-y_1)}{y_1\,(K-y_0)} \\ \beta_2 &= \log_e \frac{y_0\,(K-y_2)}{y_2\,(K-y_0)} \\ \beta_3 &= \log_e \frac{y_0\,(K-y_3)}{y_3\,(K-y_0)} \end{aligned}\right\} \quad (6)$$

by using the value of K

The first approximation to the values of the constants thus obtained is :

$$\begin{aligned} d &= +8\cdot0 \\ K &= +360\cdot15 \\ a_0 &= +5\cdot8836 \\ a_1 &= -0\cdot6224 \\ a_2 &= +0\cdot02475 \\ a_3 &= -0\cdot0004428 \end{aligned}$$

The height values are then calculated from the equation:

$$y=8\cdot0+\frac{360\cdot15}{1+e^{5\cdot8836-0\cdot6224x+0\cdot0247x^2-0\cdot00044x^3}} \quad (7)$$

and the first approximated curve is then drawn which is shown by the thick line in Fig. 1.

II. *Second approximation*

The first approximated values of the constants are then improved by finding their corrections from the normal equations which make the error of estimates a minimum.

Let the calculated values of y' from the first approximated curve

$$y'=\frac{K}{1+e^{a_0+a_1x+a_2x^2+a_3x^3}}$$

be denoted by y_1 and let their deviations from the observed values be denoted by $\triangle y$ and let $\triangle K$, $\triangle a_0$, $\triangle a_1$, $\triangle a_2$, $\triangle a_3$ be the corrections to the constants.

Then

$$\triangle y = \frac{\partial y}{\partial K} \triangle K + \frac{\partial y}{\partial a_0} \triangle a_0 + \frac{\partial y}{\partial a_1} \triangle a_1 + \frac{\partial y}{\partial a_2} \triangle a_2 + \frac{\partial y}{\partial a_3} \triangle a_3$$

$$= \frac{y_1}{K} \triangle K - \frac{y_1^2}{K} e^z \, (\triangle a_0 + x \triangle a_1 + x^2 \triangle a_2 + x^3 \triangle a_3) \qquad (8)$$

where $z = a_0 + a_1 x + a_2 x^2 + a_3 x^3$

Equation (8) can be written in a more convenient form as follows for forming the normal equations:

$$\frac{K}{y_1^2} e^{-z} \triangle y = \frac{e^{-z}}{y_1} \triangle K - (\triangle a_0 + x \triangle a_1 + x^2 \triangle a_2 + x^3 \triangle a_3)$$

The normal equations that minimize

$$\sum_x \left[\frac{K}{y_1^2} e^{-z} \triangle y - \frac{e^{-z}}{y_1} \triangle K + (\triangle a_0 + x \triangle a_1 + x^2 \triangle a_2 + x^3 \triangle a_3) \right]^2$$

are then formed.

Solving these equations the different corrections have been found. They are:

$\triangle K = -0{\cdot}003097$

$\triangle a_0 = -0{\cdot}130306$

$\triangle a_1 = +0{\cdot}0115431$

$\triangle a_2 = -0{\cdot}0003025$

$\triangle a_3 = +0{\cdot}00001882$

And the second approximating curve has the equation:

$$y = 8{\cdot}0 + \frac{360{\cdot}15}{1 + e^{5{\cdot}7533 - 0{\cdot}6109x + 0{\cdot}02445x^2 - 0{\cdot}00044x^3}} \qquad (9)$$

The final curve (9) is shown by the dotted line in Fig. 1.

The observed height values and those calculated from the initial and the final curves are given in Table 2. The standard error of estimate has decreased from 7·06 for the initial curve to 6·93 for the final curve, indicating the improvement.

4. Correlation coefficients and estimates

The yield of canes was then correlated with these growth constants. Table 3 gives these correlation coefficients along with the corresponding standard errors.

It may be seen that all the coefficients of correlation between yield and the "growth constants" are higher than their standard errors, except for a_3, indicating thereby the significance of the coefficients except the last.

Based on these coefficients of correlation, the multiple regression equation representing yield in terms of these "growth constants" is:

$$y = 0{\cdot}11372\,(K+d) + 9{\cdot}2545\,a_0 + 136{\cdot}0985\,a_1 + 1640{\cdot}7701\,a_2 - 8940{\cdot}657\,a_3 - 12{\cdot}93 \qquad (10)$$

Using this regression equation the yield of cane was estimated for all the years. These values together with the actuals and the deviation percentages are given in Table 4. The actual and the estimated yields have also been presented graphically (Fig. 2) to show the deviations at a glance.

It may be seen from the above table that the deviations of the estimated values from the actuals are within 10 per cent of the actuals in 11 cases out of 14 and in the remaining 3 cases they are within 12 per cent. The regression equation (10) has been developed by using the data upto 1959-60.

Fitting the growth curve to the 1960-61 observations, the "growth constants" for the year have been calculated and substituting these values in the regression equation (10), the estimated yield comes out as 43·2 tons/acre whereas the actual was 47·4 tons/acre. The predicted yield for the year 1960-61 is thus within a deviation of 9 per cent of the

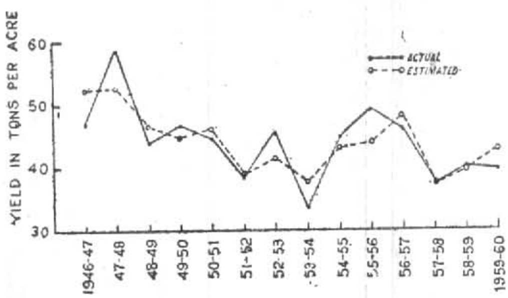

Fig. 2. Actual and estimated yields

TABLE 2

Week No. beginning with the date of first observation	Weekly mean height (cm)	Calculated value from initial curve $K=360\cdot15$, $d=8\cdot0$, $a_0=5\cdot8836$, $a_1=-\cdot6224$, $a_2=+\cdot02475$, $a_3=-\cdot0004428$	Calculated values from final curve $K=360\cdot15$, $d=8\cdot0$, $a_0=5\cdot7533$, $a_1=-\cdot6109$, $a_2=+\cdot02445$, $a_3=-\cdot0004409$
0	9·0	9·0	9.0
1	10·6	9·8	10·0
2	12·3	11·1	11·5
3	12·9	13·2	13·7
5	23·2	20·4	21·4
6	23·3	26·1	27.3
7	29·3	33·1	34·8
8	39·4	42·0	43·6
9	56·4	52·5	54·4
10	64·1	64·1	66.0
11	88·1	77·0	78·7
12	96·1	90·7	92·7
13	114·7	104·9	107·0
15	127·3	134·0	135·6
16	146·3	148·2	149·9
17	160·7	163·0	163·9
19	195·5	191·7	192·6
21	224·2	221·0	221·9
22	237·2	236·4	236·4
23	256·9	251·0	251·0
24	281·9	265·5	265·5
26	289·9	294·0	294·0
27	295·1	307·2	307·2
28	302·2	319·6	319·6
30	340·1	340·1	340·1
32	359·5	354·0	354·2
34	365·8	362·1	362·2
36	365·1	366·0	366·0
39	368·0	367·8	367·8

TABLE 3

Growth Constants	Correlation Coefficients	Standard Error
$K+d$	$+0\cdot6076$	$0\cdot2293$
a_0	$+0\cdot4811$	$0\cdot2530$
a_1	$-0\cdot4531$	$0\cdot2573$
a_2	$+0\cdot4164$	$0\cdot2625$
a_3	$-0\cdot2655$	$0\cdot2784$

TABLE 4

Year	Actual yield (Tons/acre)	Estimated yield (Tons/acre)	Deviation percentages
1946-47	47·0	52·2	11
1947-48	58·3	52·5	10
1948-49	43·9	46·5	6
1949-50	46·6	44·8	4
1950-51	44·6	46·0	3
1951-52	38·3	38·7	1
1952-53	45·4	41·4	9
1953-54	33·4	37·5	12
1954-55	45·0	43·0	4
1955-56	49·2	43·9	11
1956-57	46·0	48·1	5
1957-58	37·3	37·2	0
1958-59	40·0	39·6	1
1959-60	39·6	42·8	8

actual and establishes the usefulness of the multiple regression equation (10).

The regression equation, when improved after accumulation of a long series of data, will be useful as a working formula for predicting the yield of sugarcane two months ahead of harvest.

5. Multiple correlation coefficient

The multiple correlation coefficient R between yield and the growth constants works out to be $0 \cdot 8084$ which is significant even at 1 per cent level of probability. This means that $(0 \cdot 8084)^2$, *i.e.*, 65 per cent of the variation in yield can be accounted for by "growth constants". On the other hand the coefficient of correlation between the yield and the maximum height works out to be $0 \cdot 679$. So the maximum height accounts for 46 per cent of the variations in yield. Thus the use of "growth constants" is definitely an improvement over that of maximum height alone in forecasting yield.

6. Conclusion

From the results obtained, it can be tentatively concluded that the growth constants can be reliably used to predict the yield with a fair degree of accuracy. These "growth constants" can be computed when the heights upto the maximum stage are known. By using these constants, the yield can be forecast about 2 months in advance of the harvest. It would also be worthwhile to see if the "growth constants" themselves could be estimated from the initial portion of the growth curve (say upto the maximum rate of growth or so) in which case it would be possible to forecast the yield much ahead of the harvest.

7. Acknowledgement

The authors have great pleasure to thank Shri M. Gangopadhyaya, Director of Agricultural Meteorology, for suggesting the problem and for the keen interest he has taken in its progress.

REFERENCES

Davies, H. T. — *The Theory of Econometrics*, pp. 209-222.

Keen, B. A. 1940 *Quart. J. R. met. Soc.*, **66**.

Ramabhadran, V. K. 1950 *J. Indian Soc. agric. Stat.*, **2**, 2.