Jour. Ind. Soc. Ag. Statistics
47(3), 1995 : 253–261

Yield Forecast Using Curvilinear Study of Yield and Biometrical Characters

Rajender Parsad and R. C. Jain
Indian Agricultural Statistics Research Institute, New Delhi-110012
(Received : April, 1994)

SUMMARY

Yield forecast model using curvilinear study of yield and biometrical characters has been developed for jowar crop. The results indicate that reliable forecast of jowar yield is possible one month before harvest through this method. Optimum values of plant biometrical characters can also be obtained as a by-product of the method which are useful to plant breeders.
Key words : Curvilinear study, Graphic approximation, yield forecast.

Introduction

To introduce an element of objectivity in pre-harvest forecasts, some studies have been carried out, both in India and abroad, which use biometrical characters for pre-harvest forecasting of crop yields. This is usually done by fitting multiple regression models. In these regression models, biometrical characters have been used as explanatory variables as such or by using transformation like square root, logarithmic and reciprocal Sardana *et al* [7]; Singh *et al* [8], [9]; Jha *et al* [5], Jain *et al* [3], [4] developed regression models using plant characters data of two or more periods through growth indices/principal components as regressors. But these studies are related to linear relationship or some other assumed relationship between dependent variable (yield) and independent variables (biometrical characters) but in practice we do not know the actual relationship so as to determine the extent to which changes in the dependent variables are associated with the changes in each particular independent variables while simultaneously removing that part of variation in the dependent variable which is associated with remaining independent variables.

The actual relationship between yield and biometrical characters can be determined by successive graphic approximations in preference to fitting polynomial regression as this does not require any assumption on the shape of curves. This technique was used by Ezekiel and Fox [1]; Ramamurthi and Banerjee [6] to determine relationship between weather variables and yield.

Gangopadhyay and Sarker [2] used it for determining the relationship between crop-characteristics and yield of Sugarcane using the time series data. In the present investigation, we have used this technique to find out relationship between yield and biometrical characters and to develop forecast model for jowar (Sorghum) crop using cross-sectional data and also optimum values of various biometrical characters have been obtained which will be useful to plant breeders.

2. *Material and Methods*

The data used were collected from Tasgaon zone of Sangli district in Maharashtra State, which is situated between $16^0 75'$ N and $17^0 75'$ N latitude and between $73^0 40'$ E and $75^0 35'$ E longitude, in the years 1977-78 to 1978-79 on the hybrid jowar which has a duration of about 100 days. The sowing season is spread over the period of third week of June to third week of July, the optimum period being the last week of June to first week of July and it is harvested in October. The commonly grown varieties are CSH-1 and CSH-5. The villages, fields and plots were selected using stratified multistage random sampling design. 12 villages were selected from the zone and from each village, a simple random sample of four jowar growing fields were selected. Two sample plots of size 5 rows x 5 meters (approx. 10 sq.m.) were located randomly from each field for recording biometrical observations at fortnightly intervals beginning from one and half months after sowing. Observations were recorded at 15 days interval starting from 45 days after sowing till harvest. The periods were 6-8, 8- 10, 10-12, 12-14 weeks after sowing and at harvest. Hereafter these periods will be denoted as 1, 2, 3, 4, 5. Number of plants, plant height and number of green leaves were recorded in all five periods; length and breadth of topmost fully opened leaf in the first two periods; length and breadth of flag leaf, third leaf in last two periods and length of ear-head at harvest.

The data on 46 fields were available during 1977-78 in periods 1,2,4 and 5 whereas 40 fields data were available in period 3 instead of 48 as planned in the survey. In some cases either data were not recorded or there were incomplete recording of the data resulting in rejection of that field data. In 1978-79, 38 fields data were available in period 3 whereas 46 fields data were available in other periods.

3. *Statistical Analysis*

3.1. Scatter Graphs

Scatter graphs of yield vs. biometrical characters were plotted to find out significant biometrical characters related to yield.

3.2. Method of Successive Graphic Approximation

Method of successive graphic approximation given by Ezekeil and Fox [1] has been used to obtain suitable curvilinear relationship between yield and biometrical characters.

At each stage of approximation, the standard error of the estimate is worked out as

Standard error of estimate

$$= \left\{ \left(\sum_{i=1}^{n} (\text{observed yield} - \text{estimated yield})^2 \right) / (n-m) \right\}^{1/2}$$

Where n = number of sets of observations in the sample

m = number of parameters in the regression equation

The process is continued until the standard error of estimate show a steady value or reach a minimum.

3.2.4. Estimation of yield

To get estimate of yield the mean is adjusted by adding correction factors due to different significantly correlated crop characteristics.

The correction due to i^{th} biometrical character X_i is

$$F_i(X_i) = f_i(X_i) - M_{f(i)} \tag{3.4}$$

where $M_{f(i)}$ is the mean of the values obtained from the final curve $f_i(X_i)$, $i = 1, \ldots, p$.

The yield may now be estimated based on all crop characteristics by the equation

$$Y_0 = M + \sum_{i=1}^{P} F_i(X_i) \tag{3.5}$$

where M is mean yield.

3.2.5. Forecasting of yield

We will take the observations on relevant biometrical characters in the subsequent years upto a suitable time period and from the final regression curves and equations. Using equation(3.4) the values of F_i (X_i), i=1, . . ., p are calculated and yield forecast is done using equation (3.5).

Standard error of mean forecast is obtained as $(MSE/n)^{1/2}$. Mean squared error (MSE) $= \sum_{l=1}^{n}(Y_{ol} - Y_{fl})^2/n$ where Y_{ol} and Y_{fl} denotes the l^{th} observed and forecasted yield values respectively and n is the number of observations.

4. *Results and Conclusion*

Results based on the observations of 3rd period which corresponds to milk stage of the crop are given and forecast model is developed based on these observations. Periods 1 and 5 were excluded as 1st period was too early for developing forecast model and period 5 was too late for the purpose. Results of period 2 were not encouraging and results of period 4 were at par to period 3. Therefore, results of period 4 have also not been reported as our interest is to forecast as early as possible. Forecast model was developed on the observations based on 1977-78 and forecast of yield is done on the basis of observations during 1978-79.

The biometrical characters which were found significantly related with yield in period 3 were Plant Population (PP), Number of Green Leaves (NGL), Plant Height (PH) and Length of Flag Leaf (LFL).

The multiple regression model is as follows:

$$Y = -4.0731 + 0.0117\,PP + 1.9745\,PH + 0.0594\,NGL + 0.0803\,LFL \quad (4.1)$$

The standard error of estimate obtained from (4.1) is 0.782.

First Approximation Curve

For Plant Population

Net regression line for the plant population is obtained by substituting the mean of other biometrical characters except PP in (4.1) as

$$Y = 1.3496 + 0.0117\,PP \quad (4.2)$$

The residuals are then averaged for selected group values of PP and PP values are plotted as abscissa and residual Z_1 values as ordinates with the above net regression line as zero base *i.e.* ordinates are calculated by adding the estimated yield for particular value of PP from (4.2) and corresponding Z_1.

For line of averages, in first approximations the values of averaged residuals is given in table 1.

A free hand curve is then drawn through several group averages, as far as consistent with a smooth curve, keeping in mind the limiting conditions to the shape. The curve is first approximation to the curvilinear function $Y = f_1$ (PP). In exactly the same manner the first approximation curves for the functions $Y = f_1$ (PH) ; $Y = f_1$ (NGL) and $Y = f_1$ (LFL) are drawn.

Estimation of yield from the first approximation curve

Let us denote the first approximation curves by f_1 (PP), f_1 (PH), f_1 (NGL), f_1 (LFL), the estimates of Y_2 of Y from the first approximation curves are worked out by the equation

$$Y_2 = M + F_1 (PP) + F_1 (NGL) + F_1 (LFL) \qquad (4.3)$$

and the residuals $Z_2 = Y - Y_2$ are then computed, which are given in table 1 and the standard error of estimate is 0.841. Here M is the mean yield and $F_1(PP) = f_1 (PP) - M_{f_1pp}$ where M_{f_1pp} is the mean of all values read from first approximation curve corresponding to all values of PP. F_1 (PH), F_1 (NGL), F_1 (LFL) can similarly be explained.

Table 1. Average residuals for first, second and third approximation curves, for corresponding plant population average values.

PP values	No. of cases	Average of character values	Average of Z_1	Average of Z_2	Average of Z_3
Below 50	3	40.35	–0.57	0.12	–0.12
50–70	9	58.65	0.01	0.29	–0.03
70–90	11	79.18	0.11	–0.18	0.01
90–110	7	99.38	0.07	–0.01	–0.15
110–150	7	135.21	0.30	–0.07	0.25
150 and above	3	155.52	–0.80	0.11	–0.31

Second Approximation Curve

Now Z_2 values are averaged corresponding to suitable values of PP and the curves are plotted as deviated from the regression curves and free hand smooth curve is drawn as before.

Average values of Z_2 are given in table 1. Similarly the second approximation curves for other biometrical characters are drawn.

Now as in the first approximation curves the estimates of Y (Y_3) are obtained and $Z_3 = Y - Y_3$ are calculated which are given in table 1 and standard error of estimate is 0.782.

Third Approximation Curve

These are drawn with ordinates as average Z_3 values given in table 1 and $Z_4 = Y - Y_4$ are obtained and standard error of estimate reduces to 0.776 at the third approximation curve. As reduction in the standard error of estimate is very small, so we stop here.

These third approximation curves depicts the true relationship between yield and biometrical characters as shown in appendix.

From these graphs we see that optimum values of biometrical characters for maximum yield are given in table 2.

Table 2. Optimum values of crop characteristics.

Biometrical Characters	Optimum Values
Plant Population	130.0 – 140.0
Plant Height	1.10 – 1.20 m
Number of Green Leaves	9.5 – 10.5
Length of Flag Leaf	44.0 – 46.0 cm

Mean forecast of yield of jowar crop for year 1978-79 works out to 2.27 Kg/plot as against the actual yield 2.50 Kg/plot. Mean squared error of forecast is 0.803 and standard error of mean forecast is 0.13 Kg./plot which is about 5% of mean. On this basis, we recommend the 3rd period (10-12 weeks after sowing) for forecasting of hybrid jowar yield. Thus, reliable forecast yield of jowar can be obtained one month before harvest through this method.

ACKNOWLEDGEMENTS

Authors are thankful to the referees for their valuable comments which helped in improvement of the manuscript.

REFERENCES

[1] Ezekiel,M. and Fox, K.A., 1959. Method of correlation and regression analysis. John Wiley and Sons, New York.

[2] Gangopadhyay, M. and Sarker, R.P., 1964. A curvilinear study of yield with reference to crop characteristics - Sugarcane. *Indian J. Met. Geophy.*, **15**(2), 201-214.

[3] Jain, R.C., Sridharan, H. and Agarwal, R., 1984. Principal component technique for forecasting of sorghum yield. *Indian J. Agric. Sc.,* **54**(6), 467-470.

[4] Jain, R.C., Jha, M.P. and Agarwal, R., 1985. Use of growth indices in yield forecast. *Biometrical Journal,* **27**(4), 435-439.

[5] Jha, M.P., Jain, R.C. and Singh, D., 1981. Pre-harvest forecasting of sugarcane yield. *Indian J. Agric. Sc.,* **51**(11), 757-761.

[6] Ramamurthi, K.S. and Bannerjee, T.R., 1966. On the influence of weather on wheat yield at Dharwar. *Indian J. Met. Geophy.* **17**(4), 601-606.

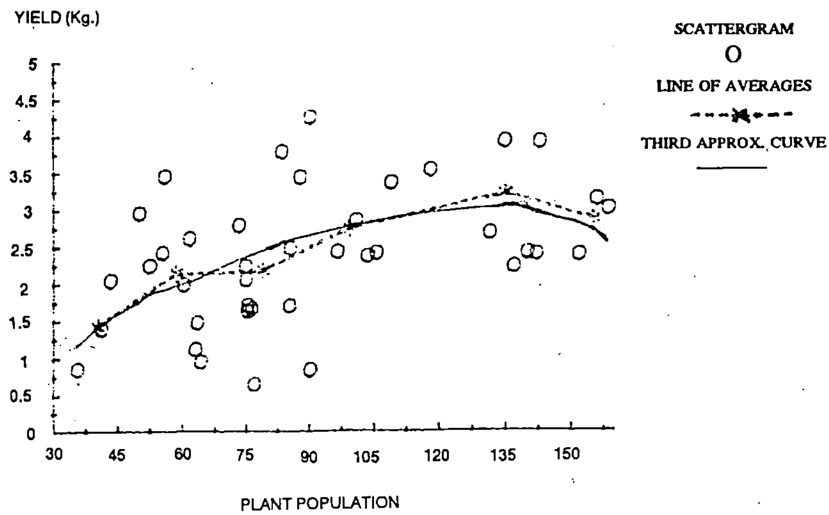
[7] Sardana, M.G., Khosla, R.K., Ohri, N.K. and Mitra, P.C., 1972. Pre-harvest forecasting of yield of jute. *Jute Bull.* **35**(7 and 8), 1-5.

[8] Singh, D., Singh, H.P., Singh, P. and Jha, M.P., 1976. Pre- harvest forecasting of wheat yield. *Indian J. Agric. Sc.*, **46**(10), 445-450.

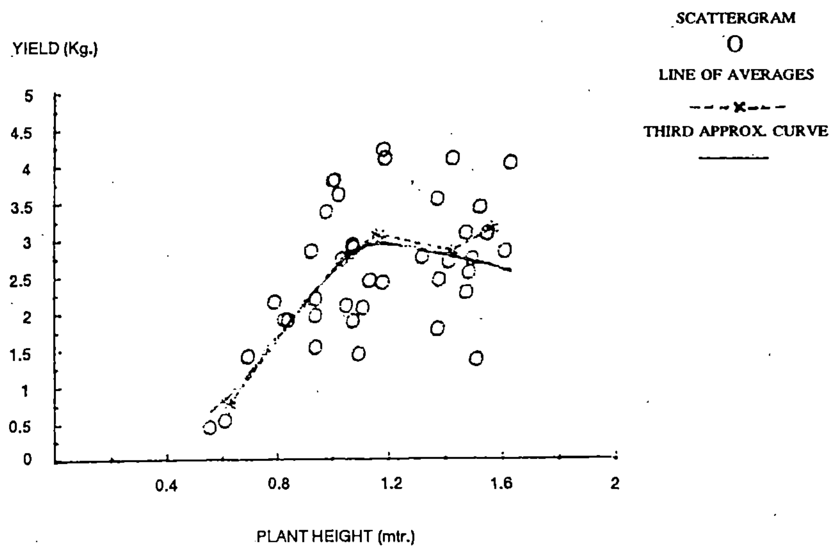
[9] Singh, D., Singh, H.P., Singh, P. and Jha, M.P., 1979. A study of pre-harvest forecasting of yield of jute. *Indian J. Agric. Res.*, **13**(3), 167-169.

APPENDIX

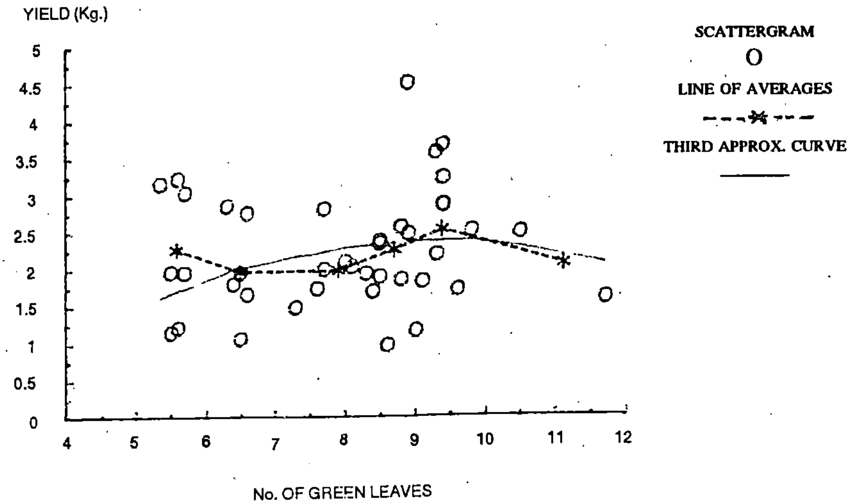
PLANT POPULATION Vs YIELD
THIRD APPROXIMATION



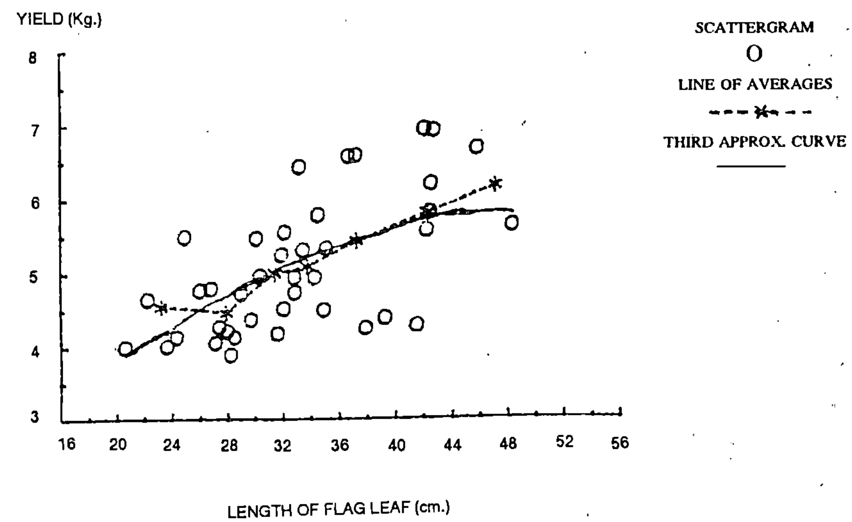
PLANT HEIGHT Vs YIELD
THIRD APPROXIMATION



NO. OF GREEN LEAVES Vs YIELD
THIRD APPROXIMATION



LENGTH OF FLAG LEAVES Vs YIELD
THIRD APPROXIMATION



Jour. Ind. Soc. Ag. Statistics
47(3), 1995 : 262–272

Some Composite Estimators for Small Area Estimation

P.S. Pandey and O.P. Kathuria*
Indian Institute of Pulses Research, Kanpur
(Received : April, 1994)

SUMMARY

Model-based composite estimators using the models given by Holt *et al* [1] have been obtained for small area estimation. The composite estimators so obtained have been empirically investigated through an illustration utilising data of sample survey conducted for estimation of milk yield of cows in different districts of Himachal Pradesh.

Key words : Composite estimators, BLUE, Synthetic estimators

Introduction

It has generally been observed that sample surveys are not able to provide estimates with adequate precision for small areas (domains) due to smallness of sample size from these domains. Alternative estimation procedures based on the "borrowing strength" from other related small areas have been proposed in the literature. The estimators based on these procedures use models either explicitly or implicitly that "connect" the small areas through supplementary data (e.g. census, administrative records etc.) Royall [3] introduced prediction approach to small area estimation. Holt *et al* [1] have discussed various models for small areas. Subsequently, several papers (Laake, [2], Sarndal [4], [5]) considered estimators using both design and model based principles. Schaible [6] considered weighted combination of synthetic and direct estimators and pointed out that exact expressions for optimal weight and variance are difficult to obtain. In this paper model-based composite estimators using the models given by Holt *et al* [1] have been obtained. The composite estimators so obtained have been empirically investigated.

2. *Notations*

Consider a finite population of size N distributed into H mutually exclusive small areas, labeled i = 1, 2, ... , H for which estimates are desired. Within each small area, units are further classified into G identifiable groups, labeled, j = 1, 2, . . . , G. The groups may be planned strata for which reliable estimates

* Presently at Markerere University, Kampala, Uganda.