CHAPTER - VII

FACTORS CONTRIBUTING INTER-FARM YIELD VARIATION OF TOBACCO

The cost price and input-output relations have been examened in Chapter-V at aggregated district and village level based on respective averages. The cost price relation at disaggregated farm level has been studied in Chapter-VI. The Chapter-VI purpots to examine as to whether the prices of tobacco formed in the open market be able to protect the interests of most of the producers in view of observed glaring inter-farm variability in cost per unit of produce. Disaggregated farm level input-output relation is yet to be analysed with a view to identifying factors contributing inter-farm variation of yield which in turn may be useful to the farmers in the decision making process of factor use in order to bring about cost price relation in his favour. The present chapter aims at finding out major factors and to estimate their contribution towards inter-farm yield variation of tobacco. Hundred and twenty same sample farmers for both the period have been chosen for the purpose of the present analysis.

Specification of Yield Determining Factors

From technical points of the physical and material inputs like level of human labour days used, the part of human labour used for topping and desuckering, and for weeding and gap filling, plough-days for preparatory tillage operation, level of N, P and K applied, frequency of letting irrigation, quantity of seed used, and quantity of plant protection chemicals applied are considered as major physical and material inputs influencing yield rate of tobacco. Apart from aforesaid physical and material factors, some qualitative factors like soil texture, fertility status of the crop land, time of transplanting, spacing in between row and plants, time and method of application of fertilizers, time of letting irrigation, crop variety grown also influence the yield rate.

The quantity of seed sown on the seed bed for a particular area to be transplanted appears to vary between individual farms but it is very different and full of arbitrariness to calculate the part thereof used for transplanting to the main field because of varying mortality rate of seedling on the seed bed. Moreover, all sample farmers have reported to use farm produced seed and not to bother about seed rate because of necessitating meagre amount of seed for transplanting a hectare of land. No accountable difference is noted as reported by the sample farmers in respect of spacing. Considering all these, the factor seed rate may be kept aside in explaining inter-farm yield variability. In view of nonavailability of soil testing report for the individual farms the factor fertility status could not be taken into consideration in the present analysis. The range of time of transplanting as reported by the selected farmers is recorded to be quite low and hence the factor time of transplanting may plausibly be assumed away for the present purpose. Time and method of fertilizer application depending upon the growth stage of crop are reported to be same, by and large, followed by the individual farmers. Ten crop varieties of local in origin are recorded to be grown by the selected farmers. But significant difference in their relative efficacy in growth performance is not reported by the farmers.

Aforementioned reported information keeping in view the following factors have been taken into consideration in explaining observed inter-farm variability of yield rate of tobacco.

- x_1 : Level of N (nitrogen) used;
- x_2 : Level of P (P_2O_5) used;
- x_3 : Level of K (K_2O) used;
- x_4 : total human labour engaged;
- x_5 : labour engaged for topping and desuckerring operation;
- x_6 : labour engaged for weeding and gap filling operation;
- x_7 : number of bullock-pair days employed;
- x_8 : number of letting irrigation;
- x_9 : Cost of plant protection chemicals incurred; and
- x_{10} : Soil texture (Sandyloam = 3, Loam = 2, Sandy = 1, Clay-loam = 0).

Analytical Tool

In contrast to estimating mean value of dependent variable (yield) on the basis of the known values of predictor variables (factors) and to work out fixed parameters of predictor variables (regression coefficients), the present purpose is to classify the dependent variable according to its value and find out the power of the individual predictor variables in classifying the dependent variable. Keeping the objective in view, the multiple linear discriminant analysis has been chosen as the analytical tool for the analysis inter-farm yield variation.

A multiple discriminant function is defined as a linear combination of a set of preditor variables which can effectively discriminate the farms belonging to two different yield groups (here two group problem has been considered). The discriminant function used for the present analysis is as in the following.

$$Z = \sum_{i=1}^{P} lixi$$

Where,

li = coefficient discriminant function of the i th predictor variable representing power of discrimination of i th predictor variable;

xi = i th predictor variable;

P = number of predictor variables (factor);

i = 1, 2, ... p;

Z = Total discriminant scores for two yield groups of farms.

The method seeks to find out the values of co-efficients (lps) such that the squared difference between the mean Z-score for the one group and the mean Z-score for the other group is as large as possible in relation to the variation of Z-scores within the groups. The determination of the value of Co-efficients (lps) necessitates the solution of the following ten equations shown in matrix notation

$$SL = D$$

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots\cdots & S_{1P} \\ S_{21} & S_{22} & \cdots\cdots & S_{2}P \\ SP_{1} & SP_{2} & \cdots\cdots & SPP \end{bmatrix}, \quad L = \begin{bmatrix} 1_1 \\ 1_2 \\ 1_p \end{bmatrix}, \quad D = \begin{bmatrix} d_1 \\ d_2 \\ d_p \end{bmatrix}$$

where,

P = 10 (x_1, x_2, ..., x_{10} specified earlier);
L_{px1} = vector of the coefficients of the discriminant function;
S_{pxp} = pooled dispersion matrix; and
D_{px1} = vector of the elements representing differences between the means of the two groups.

The discriminant function thus obtained is subjected to test of statistical significance in order to examine whether the variables (factors) considered together are effectively discriminating the farms belonging to two yield groups-high-yielding and low-yielding groups. The high-yield and low yielding groups have been formed respectively by the farms having yield above average and those having yield below the average. Two groups thus formed are furnished below :

Period	HYG	LYG	Total
I	62	58	120
II	57	63	120

The Mahalanobis D^2 Statistic has been used to measure the distance between the two groups. In terms of D^2 the statistic can be transformed into F statistic of the following form in order to examine whether the two groups are different from each other.

$$F = \frac{N_1 N_2 (N_1 + N_2 - P - 1)}{P (N_1 + N_2) (N_1 + N_2 - 2)} D^2$$

Where,

$$D^2 = \sum_{i=1}^{p} \sum_{k=1}^{p} cik\ di\ dk = \sum_{i=1}^{p} li\ di$$

C_{ik} = (i, k)th element of the inverted matrix of S;
N_1 = number of farms belonging to high-yielding group;
N_2 = number of farms belonging to low-yielding group;

The value of observed F is compared with that of tabulated F and (P) and ($N_1 + N_2 - P - 1$) d.f. at desired level of significance (1 per cent, say).

Results and Discussions

The discriminant functions for each of two periods are presented in Table 7.1. The functions for both the period is found highly significant implying that the discriminant function formulated above has sufficiently classified the farms into two yield groups. Now the question is whether any farm (or farms) is misclassified considering the classification based on mean yield that has already been made. To work out the extent of classification the Z-values of individual farms have been calculated. The mean value of Z from individual Z-values for each of two periods has been calculated thereafter. The farm of high-yielding group showing Z-value below the mean Z-value and the farm of low yielding group showing Z-value above the mean Z-value are considered to be misclassified. The extent of misclassification following the above discrimination rule is shown by Table 7.2. A sizeable extent of misclassification particularly for period I is revealed from Table 7.2. And, therefore, it is reasonable to re-estimate the discriminant coefficients after dropping out the misclassified farms. And thus the following groups are formed to find out discriminant coefficients and function.

Period	HYG	LYG	Total
I	48	47	95
II	52	58	110

Table 7.1 : Discriminant function for Period I and Period II

Factors	Coefficients (li)	t values	R^2	R^2 (adjusted)	F Calculated	F 0.01 tabulated (10, 109) d.f	Statistical inference
			Period I				
x_1	(-)0.0475	5.604					
x_2	(+)0.0026	0.226	0.6215	0.5868	17.8965	4.00	Highly
x_3	(-)0.0012	0.113					significant
x_4	(+)0.0036	1.501					
x_5	(-)0.0298	1.952					
x_6	(-)0.0539	3.940					
x_7	(+)0.0038	1.063					
x_8	(-)0.3170	2.266					
x_9	(-)0.0002	0.125					
x_{10}	(-)0.1885	1.676					
			Period II				
x_1	(-)0.0413	4.189					
x_2	(-)0.0223	2.311	0.6211	0.5863	17.8657	4.00	Highly
x_3	(+)0.0145	1.410					significant
x_4	(-)0.0020	0.664					
x_5	(+)0.0274	1.109					
x_6	(-)0.0342	2.539					
x_7	(+)0.0054	1.229					
x_8	(-)0.1470	1.074					
x_9	(-)0.0073	4.344					
x_{10}	(-)0.1359	1.177					

Table 7.2 : Extent of Misclassification

Mean of Z-values ($\overline{Z}$)	No. of mis-classified farms in high yielding group	No. of mis-classified farms in the low-yielding group	Per cent misclassfied
	Period-I		
(-) 12.75	14	11	20.83
	Period-II		
(-) 3.73	5	5	8.33

The full classified discriminant function for each period is shown in Table 7.3. The superiority of this discriminant function over that in Table 7.1 is established while comparing the values of R^2, R^2 (adjusted) and F. All these values have increased sufficiently as compared to those for earlier one. The percentage contributions of the individual factors to total distance measured are shown in Table 7.4. Use of N (x_1), labour for topping and desuckerring (x_5) and frequency of irrigation (x_9) are found statistically significant in the first period. In the second period use of N(x_1), P(x_2), K(x_3), labour for weeding and gap filling (x_6), bullock pair-days (x_7) and plant protection chemicals (x_9) are observed to be statistically significant. It implies that the above three factors have acted upon as powerful discriminators contributing about 95 per cent of total explained between group variation in the first period as revealed by Table 7.4. In the second period the aforesaid six factors together with also account for 95 per cent of explained between group variation. It is also observed from Table 7.4 that the use of N has come up as a single most important factor in both the period singularly contributing 55 per cent in the first and 53 per cent in the second period to the total explained between group variability. Thus, observed inter-farm variation in the level of N in association with other factors has contributed more than 50% to the inter-farm yield variability. Therefore, one should accord top most priority to the prudent application of nitrogen fertilizers or manures with a view to maximizing economic return out of cultivation. In regard to other important

Table 7.3 : Full classified discriminant function for Period I and Period II

Factors	Coefficients (li)	t values	R^2	R^2 (adjusted)	F Calculated	F 0.01	Statistical inference
			Period I				
x_1	(-)0.0556	5.274					
x_2	(+)0.0109	0.615	0.7490	0.7191	25.069	4.02 (10,84) D.F.	Highly Significant
x_3	(-)0.0042	0.318					
x_4	(-)0.0013	0.378					
x_5	(-)0.1163	4.971					
x_6	(-)0.0188	1.009					
x_7	(+)0.0001	0.016					
x_8	(-)0.4131	2.392					
x_9	(+)0.0006	0.285					
x_{10}	(-)0.1115	0.861					
			Period II				
x_1	(-)0.0514	5.421					
x_2	(-)0.0207	3.047	0.7406	0.7144	28.264	4.01 (10,99) D.F.	Highly Significant
x_3	(+)0.0242	2.465					
x_4	(+)0.0017	0.545					
x_5	(-)0.0060	0.225					
x_6	(-)0.0334	2.614					
x_7	(+)0.0089	2.045					
x_8	(-)0.1248	0.995					
x_9	(-)0.0089	5.787					
x_{10}	(-)0.1842	1.756					

Table 7.4 : Percentage contribution of individual factors to the total distance measured

Factors	Period I				Period II			
	Coefficients (li)	Mean difference (di)	Coefficient x Mean difference (li x di)	Per cent contribution	Coefficients (li)	Mean difference (di)	Coefficient x Mean difference (li x di)	Per cent contribution
1. Level of N used (x_1)	(-)0.0556** (5.274)	(-)31.26** (10.60)	(+)1.7380	55.30	(-)0.0514** (5.421)	(-)25.83** (10.31)	(+)1.3276	53.33
2. Level of P used (x_2)	(+)0.0109 (0.615)	(-)17.25** (6.07)	(-)0.1880	(-)5.98	(-)0.0207** (3.047)	(-)19.06** (8.47)	(+)0.3945	15.85
3. Level of K used (x_3)	(-)0.0042 (0.318)	(-)14.95** (5.23)	(+)0.0628	2.00	(+)0.0242* (2.465)	(-)13.20** (4.97)	(-)0.3194	(-)12.83
4. Total human labour (x_4)	(-)0.0013 (0.378)	(-)42.44** (5.27)	(+)0.0552	1.75	(+)0.0017 (0.545)	(-)29.51** (4.13)	(-)0.0502	(-)2.02
5. Labour for topping and desuckerring (x_5)	(-)0.1163** (4.971)	(-) 7.84** (7.530)	(+)0.9118	29.01	(-)0.0060 (0.225)	(-) 1.32 (1.53)	(+)0.0079	0.32
6. Labour for weeding and gap filling (x_6)	(-)0.0188 (1.009)	(-)10.51** (7.45)	(+)0.1976	6.30	(-)0.0334** (2.614)	(-) 5.95** (3.52)	(+)0.1987	7.98
7. No. of bullock pair days (x_7)	(+)0.0001 (0.016)	(-)12.29* (2.01)	(-)0.0012	(-)0.04	(+)0.0089* (2.045)	(+)4.17 (0.98)	(+)0.0371	1.49
8. No. of irrigation (x_8)	(-)0.4131* (2.392)	(-) 0.81** (5.47)	(+)0.3346	10.65	(-)0.1248 (0.995)	(-)0.64** (4.38)	(+)0.0799	3.22
9. Cost of p.p. chemicals (x_9)	(+)0.0006 (0.285)	(+) 1.03 (0.08)	(+)0.0006	0.02	(-)0.00089** (5.787)	(-)82.46** (9.05)	(+)0.7339	29.48
10. Soil texture (x_{10})	(-)0.1115 (0.861)	(-) 0.28 (1.46)	(+)0.0312	0.99	(-)0.1842 (1.756)	(-)0.43* (2.50)	(+)0.0792	3.18
Total			3.1426	100.00			2.4892	100.00

*stands for statistically significant at 5% level
** stands for statistically significant at 1% level
Figures in the parenthesis indicatesthe respective t-value

Table 7.5 : Percentage contribution of the important factors to the total distance measured

Factors	Period -I (Step 7)					Period-II (Step 3)				
	Coefficent (li)	F^2 (adjusted)	Mean difference (di)	Coefficient X Mean difference (lixdi)	Per cent contribution	Cofficient (li)	R^2 (adjusted)	Mean difference (di)	Coefficient X Mean difference (lixdi)	Per cent contribution
1. Level of N used (x_1)	(-)0.0588** (7.409)	0.7328	(-)31.26** (10.60)	(+)1.8381	55.22	(-)0.0514** (5.480)	0.7192	(-)25.83** (10.31)	(+)1.3276	50.26
2. Level of P used (x_2)	- -		- -	- -	- -	(-)0.0277** (3.176)		(-)19.06** (8.47)	(+)0.5280	19.99
3. Level of K used (x_3)	- -		- -	- -	- -	(+)0.0239* (2.487)		(-)13.20 (4.97)	(-)0.3155	-11.95
4. Labour for topping and desuckerring (x_5)	(-)0.1229** (5.974)		(-)7.84** (7.53)	(+)0.9635	28.94	- -		- -	- -	- -
5. Labour for weeding and gap filling (x_6)	(-)0.0176 (1.112)		(-)10.51 (7.45)	(+)0.1850	5.56	(-)0.0306** (2.946)		(-)5.95** (3.52)	(+)0.1821	6.89
6. No. of bullock pair days (x_7)	- -	- -				(+)0.0088* (2.037)		(+)4.17 (0.98)	(+)0.0367	1.39
7. No. of irrigation (x_8)	(-)0.4226** (2.732)		(-)0.81** (5.47)	(+)0.3423	10.28	(-)0.1324 (1.082)		(-)0.64** (4.38)	(+)0.0847	3.21
8. Cost of p.p. chemicals (x_9)	- -		- -		- -	(-)0.0087** (5.910)		(-)82.46** (9.05)	(+)0.7174	27.16
9. Soil texture (x_{10})	- -		- -		- -	(-)0.1865 (1.804)		(-)0.43* (2.50)	(+)0.0802	3.05
Total				3.3289	100.00				2.6412	100.00

*stands for statistically significant at 5% level

**stands for statistically significant at 1% level

Figure in the parenthesis indicates the respective t-value factor contributing inter-farm yield variation the relative importance of the factors is noted to vary widely between two periods. Labour use for topping and desuckerring, frequency of irrigation, and labour used for weeding and gap filling are noticed to be the three important discriminating factors next to N in order of importance in the first period. While in the second period the factors namely, plant protection chemicals, P, and labour used for weeding and gap filling are found to be the three important discriminators. It is worthwhile to note here that the input-output data have been collected from the same farmers in two points of time. Alongside a technological improvement in the form of yield augmentation with input rationalization is visualised in the second over the first period which has been discussed in Chapter-V. Keeping these two aspects in mind, one may plausibly lead to the contention that the technological change has brought about a change in weight structure of yield discriminating factors. And therefore, the factors like N, plant protection chemicals, P, and labour for weeding and gap filling in conjunction with other factors should be considered as key factors influencing yield of tobacco under present set of production technology.

Now the question of dropping out the discriminating factors that are found redundants and to reformulate the function based on important discriminators comes to the fore. The step down results reveals that the value of R^2 (adjusted) has been continuously increasing up to step 7 in case of the first period and up to step three for the second period. In view of this, the results of the respective step for the two periods have been chosen and summerised in Table 7.5. A little over 97 per cent of explained between group variability is attributable to the variability of application of N, P, and plant protection chemicals in the second periods as evinced by Table 7.5. Thus, programme aiming at enhancing yield one should emphasize on the use of these three key inputs. The economic decisions in respects of level to be applied, however, depends upon the response function and price of the input under consideration and the expected price of the produce.