

(6) Correlational analysis did not reveal any significant influence of rainfall and various other factors on the yield per bigha.

(7) The correlation between yield per bigha and the number of fever deaths in the Bolpur Thana was -0.37 (for $n=24$ years) and was possibly just significant. When many people have died of fever in a particular year we know that many more must have suffered from fever for a long time. This would withdraw many workers from the field and thus exert a harmful effect on production. The connexion between fever deaths and yield is therefore quite plausible.

(Received December, 1933).

EDITORIAL NOTE

It will be seen from Table 1 that the size and number of plots as well as the total area for which records are available are fluctuating considerably from year to year. For any particular year we can obtain the 'average yield' either by dividing the total yield by the total acreage (given in Table 1), or from the frequency distributions of yield (given in Table 3); the two values will in general be different, since the size of the plots is not constant. From either set of average yields for particular years we can then find the trend (a) by giving the same weight to the averages, or (b) by giving different weights (according to the total area, or the total number of plots as the case may be). We shall therefore obtain four distinct solutions corresponding to:—

- (1) Averages based on total areas without weighting.
- (2) Averages based on total areas with weighting.
- (3) Averages based on frequency distribution without weighting.
- (4) Averages based on frequency distribution with weighting.

The significance of each of the above solutions will naturally be different. Let us consider the chief factors which cause variations in the yield.

(a) The climatic factors and conditions of cultivation change from year to year and cause variations in the yield even in the same plot.

(b) The fertility of the soil varies from plot to plot, so that the yields are different in different plots even in the same year.

(c) Records of yield in any particular year are not available for the whole but only for a portion of the total area under cultivation in the village, and as this portion varies from year to year (not necessarily in a random manner), certain further fluctuations are imposed on the estimates of the yield.

(d) We also have the errors of measurement, and finally the random fluctuations due to a variety of small causes.

If the primary data had been presented in complete detail with the size and yield of each plot in each year, we could make an attempt to assess directly the relative importance of the different factors of variation. Unfortunately the condensed form of the data renders this impossible, and we are obliged to have recourse to certain general considerations.

TABLE N 1. YIELDS OF RICE IN MAUNDS PER BIGHA.

No.	Estimate of Yield based on	Weights	YIELD IN MAUNDS PER BIGHA			STANDARD DEVIATION IN MAUNDS PER BIGHA		
			Owner-cultivated	Share-system	Total	Owner-cultivated	Share	Total
(1)	Total area ...	Unweighted ..	6'40 ± '26	4'37 ± '18	5'58 ± '22	2'40 ± '19	1'68 ± '18	2'01 ± '16
(2)	Total area ...	Weighted ...	6'52 ± '30	4'27 ± '17	5'40 ± '19	2'74 ± '21	1'58 ± '12	1'76 ± '14
(3)	Number of plots ...	Unweighted ...	7'32 ± '05	5'34 ± '04	6'56 ± '08	2'14 ± '08	1'30 ± '08	1'82 ± '02
(4)	Number of plots ...	Weighted ...	7'74 ± '07	5'22 ± '07	6'79 ± '05	3'00 ± '04	2'48 ± '05	3'05 ± '04

TABLE N 2. YIELDS OF RICE IN LBS. PER ACRE.

No.	Estimate of Yield based on	Weights	YIELD IN lb PER ACRE			STANDARD DEVIATION IN lb PER ACRE		
			Owner-cultivated	Share system	Total	Owner-cultivated	Share	Total
(1)	Total area ..	Unweighted ...	1593 ± 65	1058 ± 46	1389 ± 55	597 ± 46	418 ± 32	500 ± 39
(2)	Total area ...	Weighted ...	1623 ± 75	1063 ± 42	1344 ± 48	682 ± 53	381 ± 29	433 ± 34
(3)	Number of plots ...	Unweighted ...	1822 ± 12	1329 ± 10	1633 ± 7	533 ± 7	324 ± 7	453 ± 5
(4)	Number of plots ...	Weighted ...	1927 ± 17	1299 ± 17	1690 ± 12	747 ± 10	605 ± 12	759 ± 10

THIRTY-EIGHT YEARS OF RICE YIELDS

It will be easily seen that the number of plots will control factor (d), but in the absence of any assurance that the plots constitute a random sample it is doubtful how much value can be attached to mere replication. In fact, in case the samples are not random, weighting by the number of plots is likely to make the estimates of average yield

The total area for which records are available on the other hand will control factors (b) and (c). When the total area for which records are available is large it will include soils of different grades and will thus give a reliable estimate of the average yield; on the other hand, when the total area is small the estimate may not be fully representative. On this view, weighting according to area would be desirable.

Finally in case the seasonal factor (a) is the dominant one (in the sense that fluctuations from year to year are more important than fluctuations from field to field within the same year), then it is clear that it will not matter very much whether we work with the unweighted or the weighted averages.

In view of the interest of the question I am giving the results obtained by different methods for $n = 38$ years in Tables N1 and N2 (p: 388).

Fitting a straight line by the direct correlation method we find the following constants.

TABLE N (3)

Method of calculation		"Weights"	r = Coefficient of correlation	b = Coefficient of Regression	Ratio of Owner : Share
1	Total Area	Unweighted ...	-0'0087	-0'0016	1'46
2	Total Area	Weighted ...	-0'0098	-0'0016	1'53
3	Number of Plots	Unweighted ...	-0'0068	-0'0446	1'37
4	Number of Plots	Weighted ...	-0'0319	-0'0047	1'48

Several things are clear from the above results:—

(i) The ratio of the yield for Owner-cultivated to the yield under the Share-system comes out substantially the same (about 1.5) in all cases.

(ii) The results for 'weighted' and 'unweighted' averages are in close agreement. This suggests that the seasonal changes are probably the most important causes of variations in the yield.

(iii) The coefficients of correlation, and regression are negligibly small in all cases.

We may therefore conclude that there has not occurred any appreciable change in the soil-fertility so that the average yield has on the whole remained practically the same during the period under review. The yield under Owner-cultivation has however been about 50 per cent. higher than the yield under the Share-system.

As regards actual estimates I think on the whole it will be preferable to adopt the weighted averages based on total areas which give the following trend equation:—

$$y = 5.40 - 0.0016 (t - 1913)$$

where y = yield of paddy in maunds per bigha and t = year.

—P. C. Mahalanobis.