PLOT SIZE IN CROP CUTTING SURVEYS ON PADDY

By

D. Singh, K. S. Krishnan And P.N. Bhargava

Institute of Agricultural Research Statistics, New Delhi-12

Received in October, 1974

Random crop cutting technique has been recognized to be the most reliable method for estimating yield rates of crops. The technique consists in sampling a unit of area based on probability sampling method, harvesting the crop and obtaining its produce. A good amount of work has been done in India and in many other countries for standardizing the procedure to be adopted in yield estimation surveys conducted by various Governments. The problem of size and shape of the unit area that should be ultimately sampled for harvesting has drawn considerable attention of the Agricultural Statisticians in India and elsewhere.

Indian Statisticians (Mahalanobis-1945, Sukhatme-1946 and Panse-1954-63) have generally observed that very small cut as ultimate unit of sampling gives an over-estimation of yield rate. This over-estimation is attributed to causes such as location bias and error in demarcation of plot. Location bias appears to be more serious in cases of uneven crop stand while bias due to demarcation is generally caused by the mis-classification of plants on the border of the plot by the field investigator. In the case of uneven crop, if by chance the random cut falls in an area where there is no crop—relative probability of which is generally more in the case of small plots—the general tendency of the field investigator is to shift the random cut in such a manner that he may be able to include some plants so that he can avoid reporting zero yield. In deciding the inclusion within or exclusion from the plot the border plants, the general tendency of the field investigator is to include even those plants, major portion of whose root system are lying outside the plot boundary and hence should be excluded. This leads obviously to an over-estimation of yield rate and its effect is relatively more serious in plots of small sizes.

The plot size investigations on paddy, the principal foodgrain crop in India have been conducted recently under normal field

conditions obtaining in the country. The first study was conducted during 1962-63 in Orissa State and the second during 1963-64 in Andhra Pradesh. Their results are reported in the present paper.

Material Used

A random crop cutting survey for estimation of yield rate and total production of paddy crop in Orissa State is conducted annually by the State Bureau of Statistics and Economics. The State comprises of 13 districts and each district has been divided into a number of sub-divisions for administrative purposes. The number of sub-divisions in a district ranged from 9 to 15. Each sub-division consists generally of about 400 to 500 villages. For the purposes of crop surveys these sub-divisions are treated as strata. In each stratum crop cutting experiments are planned on a random sample of paddy growing fields. The actual method of selection consists in selecting the specified number of villages in each stratum on probability basis and thereafter choosing in each sample village three fields randomly for the conduct of crop cutting experiments. In each selected field a circular cut of 4' radius is randomly located and its produce is harvested. Within a circular cut of 4' radius another cut of 2'-3" radius is taken. The produce from this inner cut is kept for driage. The dry grain-wet yield ratio is subsequently used for converting the wet yield obtained from 4' radius cut to dry grain.

In the course of field work during 1962-63, a field investigator responsible for the conduct of crop cutting survey was instructed to carry out in one of the three fields in each of the sample village, an additional crop-cut of size $16\frac{1}{2}' \times 16\frac{1}{2}'$. The objective of the conduct of this additional cut was to study the bias, if any, in the estimation of yield rate with the adoption of small plot. Yield data for both the types of cuts were obtained in all from 1137 fields distributed over the different districts in the state which have been made use of in the present study. The results obtained from the study are of great practical significance as the investigation was conducted under normal field conditions and the extent and type of supervision was on a scale normally practicable.

The data relating to Andhra Pradesh* used in the present study was obtained through a sample survey carried out jointly by the State Directorate of Economics and Statistics and the Directorate of

*The data utilised for the study was taken from the type II study conducted under the direction of the Technical Committee on Crop Estimates appointed by the Planning Commission, Govt. of India, under the Chairmanship of Dr. S.R. Sen.

National Sample Survey. The objective of the study was to investigate whether small cut gives an over-estimation of the yield rate of paddy. The second feature of this study was that in addition to data of small cut and large cut, total produce from the whole field was also obtained. This was considered useful as the yield rate based on whole field harvest may be considered as free from bias. The produce of whole field was dried and the weight of dry grains was recorded. It was contemplated at the time of initiation of the programme that the field work would be carried out under normal field conditions; but, due to the special features of the study, the institutions concerned evinced great interest, which resulted in an intensive supervision involving all the experiments where whole field harvest was carried out. This made the field investigators responsible for crop cutting work extra careful in the demarcation of plots and harvesting of produce. Therefore, in the case of experiments in which whole field harvest was also done the normal field conditions did not prevail as was contemplated.

The design adopted for the survey in Andhra Pradesh was multi-stage stratified random sampling. For the purpose of the survey, the State was divided into 28 strata. In each stratum four villages, and in each village two fields, were randomly chosen by giving equal probability to each field. The study was planned in 112 villages. In each selected field two cuts, one of a circular shape of radius 4′ and another of rectangular shape size 33′ × 13′-2.4″ was taken. Within the circular cut a further concentric circle of radius 2′-3″ was also demarcated. Since both the types of cuts were to be located in every sample field, it was decided to locate the circular cut first in the firsf sample field in a village followed by the location of the rectangular cut, while the order of location was inter-changed for the second sample field in the same village. In no case, overlapping was allowed.

Of the 224 fields selected for the survey, data on circular and rectangular cuts were available only in respect of 193 fields. In addition, whole field harvest was planned and conducted in 112 fields sampled from 56 out of 112 randomly selected villages.

Method of Analysis

The method of analysis followed consisted of :

(1) calculating the yield rate of dry grain of paddy in kg. per acre for each cut of different sizes,

(2) working out the differences between the yield rates of grain for different cut sizes located in each field and

(3) to carry out 't' tests at 5% and 1% levels of significance to ascertain whether the mean differences are significantly different from zero.

Recognising the fact that past investigations carried out in India and elsewhere by a number of research workers have indicated the likelihood of over-estimation of yield rate in normal crop estimation surveys when small size cuts were used and this over-estimamation becomes negligible for relatively large cut, the following hypotheses were considered in formulating the appropriate tests:

Null Hypothesis	*Alternative Hypothesis*
$H_o\,(\bar{Y}_S = \bar{Y}_L)$	$H'(\bar{Y}_S > \bar{Y}_L)$
(There is no difference in yield rate due to size and shape of cut.)	(A small cut gives an over estimate of yield rate compared to that obtained with a large size cut.)

where $\bar{Y}_S$ is the true mean yield based on small cut and $\bar{Y}_L$ that based on large cut. To test the above hypothesis single tailed 't' test is the most appropriate and uniformly most powerful test. Therefore, this test has been used in both the studies discussed in the present paper.

The number of experiments conducted in a sub-division was small so that the mean yields for different cut sizes and the differences between them were obtained at district level. The district estimates were pooled to obtain estimates for the State as a whole by weighting with the number of experiments conducted in a district. This may be justified because the total number of experiments planned was distributed over different districts roughly proportional to area under the crop. The analysis was made in respect of circular cut and square cut for which data were available. Standard errors of estimates were obtained following the procedure indicated in the Appendix.

In respect of study relating to Andhra Pradesh, the 28 strata for which data were available were grouped into three zones on the basis of rainfall and agroclimatic conditions (care was taken that the jurisdiction of a stratum fall wholly within a zone). The districts included in each zone are indicated at the foot of Table III. The analysis was done separately for each zone at the first instance; thereafter pooled estimates for the State as a whole were worked out by weighting the zonal estimates with the number of experiments.

Comparisons were made between circular cut and rectangular cut, circular cut and whole field harvest and also rectangular cut and whole field harvest. Standard errors of estimates were worked out according to the procedure indicated in the Appendix.

Results and Conclusions

The results of analysis of data relating to Orissa, is presented in Table I:

TABLE I

Estimates of mean differences in the yield rates of paddy crop within the small and large cuts and their standard errors—Orissa, 1962-63

S. No.	*Name of the district*	*No. of fields Sampled*	*Circular cut—Square cut*		*Value of 't'*	*%age over-estimation by circular cut over square cut*
			Mean difference (Kg/ac.)	*S.E. of the difference (Kg/ac.)*		
1.	Balsore	80	17.4	21.47	0.81	3.8
2.	Bolangir	70	—21.4	31.32	0.68	—4.2
3.	Cuttack	117	—51.0	17.74	—2.88**	—7.7
4.	Dhenkanal	74	51.0	22.75	2.24**	10.3
5.	Ganjam	82	2.9	28.63	0.10	0.4
6.	Kalahandi	96	—6.4	31.30	—0.20	—0.9
7.	Koonjhar	60	82.7	21.70	3.81**	16 8
8.	Koraput	134	28.2	18.44	1.53	5.1
9.	Mayurbhanj	85	73.2	23.47	3.12**	11.5
10.	Phulbani	64	35.2	36.01	0.98	5.6
11.	Puri	79	88.0	30.15	4.37**	12.4
12.	Sambalpur	138	104.2	16.95	6.15**	22.5
13.	Sundergarh	34	45.5	16.69	2.73**	9.9
	Pooled	1137	33.6	6.54	5.14**	5.8

** Indicates significance at 1% level.

It may be seen from Table I that the mean difference in yield rates (circular cut-square cut) varied from 51 kg. per acre in Cuttack district to 104.2 kg. per acre in Sambalpur district. In 10 cuts of a total of 13 districts the small circular cut over-estimated the yield rate compared to the square cut. The extent of over-estimation varied from about 1% to 23%. In 6 out of the 10 cases, the over-estimate was found to be highly significant (1% level of significance).

There was, however, a single case of a significant under-estimation due to the small circular cut compared to the large square cut (Cuttack district). Thus on the whole, it may be concluded that data collected in Orissa survey confirms the earlier findings that adoption of small circular cut in large scale crop estimation surveys is likely to lead to over-estimation of yield rates. At the State level an over-estimation of 33.6 kg. of paddy per acre was obtained compared to the square cut. This was of the order of 5.8% of the mean yield of paddy grains for the State during the year of the survey.

The results in respect of Andhra Pradesh are presented in Tables II, III and IV.

Table II gives the mean grain yield of paddy in kg. per acre for each of the three zones, as well as for the State as a whole based on the data from circular cuts, rectangular cuts and whole field harvest

TABLE II

Estimates of average yield of paddy based on different cut sizes—Andhra Pradesh 1963-64

Zone No.	*No. of fields*	*Mean yield (Kg/acre)*			*Percentage over-estimation over whole field*	
		Circular cut (4′ radius)	*Rectangular cut (33′ × 13′-2.4″)*	*Whole field*	*Circular cut (4′ radius)*	*Rectangular cut (33′×13′ -2.4″)*
1	2	3	4	5	6	7
1.	40	896.5	855.1	845.1	6.1	1.2
2.	38	893.2	795.7	842.7	6.0	—5.6
3.	34	1015 0	1008.0	946.7	7.2	6.5
Pooled	112	931.3	881.4	875.1	6.4	0.7

respectively. Percentage over-estimation of circular and rectangular cuts over whole field harvest have also been given in the same Table. The data pertains to the 112 fields for which records are available for the whole field harvest.

Table III presents for the same data from 112 fields the mean differences, their standard errors and the values of '*t*' together with their degrees of freedom (in this Table also the results are presented separately for each of the three zones and then the pooled result for the State as a whole).

TABLE III

Estimates of mean differences in the yield rates of paddy crop based on different cut sizes located in fields for which whole field harvesting was done—Andhra Pradesh, 1963-64

Zone No.	No. of fields	*Circular cut minus Rectangular cut*				*Circular cut minus whole field*				*Rectangular cut minus whole field*			
		Mean difference (Kg/acre)	*S.E. of difference (Kg/ac.)*	*Value of 't'*	*d.f.*	*Mean difference (Kg/acre)*	*S.E. of difference (Kg/ac.)*	*Value of 't'*	*d.f.*	*Mean difference (Kg/ac.)*	*S.E. of difference (Kg/ac.)*	*Value of 't'*	*d.f.*
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.	40	41.4	27.12	1.53	39	51.4	31.99	1.61	39	9.96	15.74	0.68	39
2.	38	97.4	51.17	1.90*	37	50.5	46 54	1.08	37	—46.96	48.42	—1.08	37
3.	34	7.0	53.09	0.13	33	68.2	41.84	1.06	33	61.25	21.00	3.81	33
Pooled	112	50.0	25.58	1.95*	109	56.2	24.09	2.33*	109	6.2	14.81	0.42	109

* Indicates significance at 5% level.

Note :— Zone 1 : Districts of Srikakulam, Visakhapatnam, East Godavari, West Godavari, Krishna and Guntur.

Zone 2 : Districts of Adilabad, Karimnagar, Warangal, Khamam, Nizamabad, Medak, Hyderabad, Nelgonda and Mahbubnagar.

Zone 3 : Districts of Kurnool, Anantapur, Cuddapa, Nellore and Chittor.

TABLE IV

Estimates of mean differences in the yield rates of paddy crop based on the entire data on the circular and the rectangular cuts— Andhra Pradesh, 1963-64

Zone No.	*No. of fields*	*Circular cut minus rectangular cut*			
		Mean difference (Kg/acre)	*S.E. of difference (Kg/acre)*	*Value of 't'*	*d.f.*
1.	75	11.1	20.94	0.53	74
2.	70	78.2	34.72	2.25*	69
3.	48	0.3	43.00	0.01	47
Pooled	193	32.8	18.41	1.78	190

As has been mentioned earlier, in addition to the 112 fields, yield data in respect of circular cut and rectangular cut only were available for an additional 81 fields. In order to obtain more precise comparisons between circular and rectangular cuts, the data from 193 fields were utilized and the results are presented in Table IV.

The results presented in Tables II and III indicated that the mean yield for the circular cut was higher than that obtained for the rectangular cut as well as for the whole field harvest for all the three zones as well as for the State as a whole. The over-estimation with the use of circular cut is likely to be of the order of 6% on the whole compared to the rectangular cut as well as the whole field harvest. On the other hand, the overall bias in the case of rectangular cut compared to whole field harvest was seen to be less than one per cent.

The results presented in Table IV also indicate an over-estimation of the order of 32.8 kg. per acre at the State level for the circular cut compared to the rectangular cut. The '*t*' value of 1.78 based upon the data from 193 fields was statistically significant at 5% level (single tailed).

Summary and Conclusion

Random sampling procedure has been recognised to be the most suitable method for the estimation of yield rate of agricultural crops. In large scale sample surveys of this nature, apart from the problem of distribution of number of sampling units among different stages, the choice of proper unit of sample cut is also very important. In order to determine the suitable size and shape of cut for the surveys on crop production, a number of studies have been made in India and elsewhere. Most of these studies have established that small size cuts over-estimate the yield rate. This over-estimation is due to

location as well as border bias. Recently, two studies in this connection have been made, one in Orissa State 1962-63, and the other in Andhra Pradesh 1963-64. Both the studies were carried out on paddy crop. The results of these investigations are discussed in the present paper. The results obtained from the study are of great practical significance as the investigation was conducted under normal field conditions and the extent and type of supervision was on a scale normally practicable. Sampling design adopted for both the studies was multi-stage stratified random sample. The results are based on 1137 pairs of cuts from Orissa State, and 193 pairs of cuts from Andhra Pradesh. In Andhra Pradesh, in addition to data on small and large cuts, the record of total produce from 112 selected fields were also available. Both the studies indicated that small cut leads to over-estimation. The order of over-estimation in both the cases was about 6%. Over-estimation for small circular cut as compared to large cut was statistically significant in both the cases. To sum up, the results of the two surveys indicate that even in a crop like paddy for which relatively uniform crop stand is expected within a field, there is a risk of over-estimation of yield rate by adopting a very small size cut.

Acknowledgement

The study was taken up on the initiative of Dr. G.R. Seth, former Statistical Adviser, I.C.A.R., New Delhi. We are grateful to Dr. C. Mishra, Director, Bureau of Statistics and Economics, Orissa, for making available the data of the survey carried out in 1962-63. The authors acknowledge the assistance in the analysis of data given by M/s. V.S. Rastogi and S.N. Mathur, IARS, New Delhi.

REFERENCES

Mahalanobis, P.C. (1945) : Report on the Bihar Crop Survey, Rabi Season 1943-44. Sankhya 7, 29-106.

Panse, V.G. (1954) : Estimation of crop yields. Food and Agricultural Organization of the United Nations, Rome.

Panse, V.G. (1963) : Plot size again. Journal of Indian Society of Agricultural Statistics, 15 (No. 10 & 12), 151-159.

Sukhatme, P.V. (1946) : Bias in the use of small size plots in sample survey for yield. Nature, 157, 63.

— (1947) : The Problem of plot size in large scale yield surveys. Journal of American Statistical Association 42, 297-310.

APPENDIX

Let Y_{ijk} be the yield rate based on circular cut for the i^{th} village of j^{th} stratum of k^{th} district.

Let Z_{ijk} be the yield rate based on rectangular (or square) cut for the i^{th} village of j^{th} stratum of k^{th} district.

Let the difference between the yield rates due to different types of cuts for the sample plot in the i^{th} village of j^{th} stratum of k^{th} district be denoted by d_{ijk} (i.e., $Y_{ijk}-Z_{ijk}=d_{ijk}$).

Now d_{ijk} may be assumed to be a random variable distributed normally with mean zero and a constant variance σ^2.

The mean difference for the j^{th} stratum of the k^{th} district is given by $d_{\cdot jk} = \sum_j d_{ijk}/n_{jk}$.

where n_{jk} is the number of plots in the j^{th} stratum of the k^{th} district.

The mean difference for the k^{th} district is given by $d..k = \left(\sum_j n_{jk}\, d_{\cdot jk}\right)/n_k$ where n_k is the number of plots in the k^{th} district.

The mean difference over all the districts is given by

$$d... = \left(\sum_k n_k \,.\, d..k\right) \Big/ \sum_k n_k$$

The mean square deviation for the j^{th} stratum of the k^{th} district is given by

$$S^2_{jk} = (d_{ijk} - d_{\cdot jk})^2/(n_{jk}{}^{-1})$$

The mean square deviation for the k^{th} district is given by

$$S_k^2 = \frac{\sum_j (n_k{}^{-1})\, S^2_{jk}}{\sum_j (n_k{}^{-1})};$$

$$t(n_k{}^{-1}) = \frac{d..k \cdot n_k}{S_k}$$

which is used for testing the hypothesis mentioned earlier.