

# AN INTERIM REPORT ON THE CROP-CUTTING SURVEY FOR ESTIMATING THE OUT-TURN OF WHEAT IN THE PUNJAB (1943-44)

BY DR. P. V. SUKHATME, PH.D., D.Sc., F.A.Sc.

Received December 9, 1944

THE fact that yield statistics of food grains crops in India are extremely unsatisfactory, and the principle that only crop-cutting experiments carried out on randomly selected plots give reliable estimates of the average yield per acre have been recognised by those concerned with the estimation of crop yields for a long time. The first authoritative recommendation that crop-cutting experiments must be done in randomly selected villages and in randomly selected fields came from the Board of Agriculture in the year 1919. The earliest experiments based on the principle of random sampling were carried out by Hubback (later Sir John Hubback) in the years 1923-25. Hubback's methods were used by Deshmukh (now Sir C. D. Deshmukh) in the C.P. in 1928-30. But apart from these attempts, no Provincial Government with the exception of Madras appears to have taken any measures to give effect to the recommendation of the Board even though this recommendation had been repeated by the Board and other high officials since the Board's first meeting in 1919. The two probable reasons for this situation seem to be (1) unwillingness to make any serious departure from the traditional departmental routine by the introduction of innovations, and (2) unwillingness to incur permanently any heavy additional expenditure involved in augmenting existing staff which it was thought would be necessary in adopting the method of random sampling.

2. The whole question of improving the yield statistics was considered at a meeting of the Inter-Departmental Committee of the Government of India which recommended that I should prepare a scheme keeping in view that only the minimum essential changes are made in the current procedure of conducting crop-cutting experiments, and that further the technique should be so simple that it can be worked by the existing staff, who have been already conducting these experiments, without any heavy additional expenditure. Accordingly I prepared a scheme for the sample survey which was approved by the Committee at its next meeting. The original scheme will be reproduced in the fuller report. In this interim report I shall give

only a brief account of the organisation which was set up in the Punjab and the plan of work actually adopted.

3. The survey was organised under the administrative control of the Director of Agriculture, assisted by his departmental staff comprising of the Assistant Director of Agriculture, the Deputy Directors of Agriculture in charge of circles, the Extra Assistant Directors of Agriculture in charge of districts and the Agricultural Assistants in charge of tehsils. A Statistician was also attached to the survey with his headquarters at Lahore for assisting the field staff in their work.

4. The survey was organised in 27 out of the 29 districts of the Province. The Lyallpur district could not be included as some special investigation connected with the technique of the survey was in progress. The hilly district of Simla was left out because the area under wheat in the district is negligible. In what follows in this report, the word 'Province' denotes the 27 districts covered by this survey.

5. The plan of sampling was based primarily on the results of previous sample survey conducted by the I.C.C.C. on cotton in the district of Akola. Briefly the conclusion reached in the Akola survey was that a plan of sampling by tehsils which includes approximately 2 per cent. of the crop-growing villages in the Province, 3 fields in each village and one plot in each field is practicable and adequate to furnish (a) a reliable estimate of the average yield for the Province as a whole; and (b) reliable quinquennial estimates of 'normals' in each district.

6. A total of 748 villages, representing approximately two per cent. of the total number of villages in 27 districts was selected under the plan. These were divided among the 27 districts approximately in proportion to the area under wheat in each district, but within a district the number was divided equally between the tehsils.

7. The villages within a tehsil were selected at random. The selection was made with the help of printed random numbers to ensure equal chance of inclusion for every village in the tehsil. The selection was done personally by me at the centre, since I had anticipated that considerations of distance, of lack of communications and other difficulties might weigh with the provincial staff and be allowed to interfere with the random character of the sample.

8. Within a selected village three random fields from amongst all the wheat-growing fields, pure and mixed, were chosen and within a selected field, a plot measuring 66' × 33' was located at random. The selection of fields

and plots was done by the Agricultural Assistants but was supervised by the Extra Assistant Directors of Agriculture, and the Deputy Directors of Agriculture. The selection was so made that every wheat-growing field in the selected village got an equal chance of being included in the sample. The plot in a selected field was also marked at random in accordance with the instructions drawn for the purpose. All the operations of harvesting, threshing, winnowing and weighing were carried out in the presence of the Agricultural Assistants on days previously fixed in consultation with the owners. Although the fields were selected in advance to ensure that all fields in the village got an equal chance of being sampled, the plots were marked on the dates on which the produce was harvested.

Under the scheme villages and fields with smaller area under wheat will proportionately get larger representation in the sample. This, however, is unavoidable since there is no other practical method of selecting a random sample of plots. Even if there is association between the yield and area in the village (or field), the resulting bias in yield estimate can be corrected by weighting each village (or field) by its area when calculating the average.

9. It is known that the estimate of the average yield for a tehsil attains the maximum precision when the proposed number of experiments is so distributed that one experiment each is conducted in a different village of the tehsil. A word is, therefore, necessary to explain the decision to have three experiments in a village, and one experiment in a field. There are clearly three sources of variation within a tehsil affecting its average yield: (a) the variation between villages, (b) the variation between fields in a village and (c) the variation between plots in a field. Of these (a) and (b) are generally found to be appreciably larger than (c) and explain why one plot was fixed in each field. As between (a) and (b) no appreciable difference is noticed. The distribution of experiments while depending on the ratio of these variations also depends on other considerations, such as the staff available, the time required to travel from one village to another and the necessity of completing the work in the tehsil within the period available for harvesting. As only one agricultural assistant was available per tehsil and as he can, at best, be expected to visit between 6 to 8 villages during the harvesting period and as further the variability from field to field was pronounced the consideration of spreading our experiments over as large a number of villages as possible and yet of conducting all of these within the available time led us to adopt the scheme as outlined above.

10. Departmental crop-cutting experiments, everywhere in India, are done on large plots of the order of 1/10th of an acre size. On the other

hand previous sample surveys in India, in England and U.S.A. have been mostly conducted, on small size plots. The factors which led me to the choice of large plots are:—

- (i) The coefficient of variation of plot yields decreases as the plot size increases, but within the range of large size plots the choice of a larger plot does not appreciably reduce the standard error of the estimated average yield.
- (ii) The estimated average yield attains the maximum precision when the proposed number of plots in a village is so distributed that one plot each is allotted to a different field in the village. One plot per field of large size is, therefore, statistically superior to several small size plots in a field.
- (iii) The results of the previous surveys conducted by the Indian Central Cotton Committee and by the Imperial Council of Agricultural Research show that the internal variation of a field is a very minor component of the gross variability of the final estimate and that, therefore, relative to the effect of number of experiments and their distribution, the effect of plot size on the precision of the average yield is negligible.
- (iv) The choice of large plot is in keeping with the principle that the surveys should be conducted with as little modifications in current practice as possible.
- (v) With the unevenly sown crops in India, the use of very small size plots in the hands of the existing staff who are accustomed to large size plots would lead to large errors in the location and measurement of plots. Moreover, there is a risk that considerable losses will occur in the handling of small produce and in its weighing unless arrangements are possible for processing the produce at a central place under expert supervision and with the use of sensitive balances. The large size plot is mostly free from these limitations.

Simplicity of procedure was obviously important in recommending the use of the sample survey technique by departmental agencies. The necessity of accommodating the plot within the fields without having to reject too many fields and the necessity of completing harvesting and the connected operations during the course of the day lest operations may remain incomplete and the produce may have to be left over-night without adequate supervision suggested that the plot size of 1/20th of an acre measuring 66' × 33,' would be suitable pending further investigations on the point.

There is an inherent difficulty in sampling from a finite field to which my attention has been drawn by Fisher. Whatever be the size of plot, the central portion of the field is relatively over-sampled as compared with the wide surrounding areas. The bias in sampling becomes smaller with smaller size plot, but is not completely removed. The remedy lies in introducing a correction factor such that it would equalize the chance for every portion of the field being included in the plot. If  $p$  is the probability that a sample plot occupies a particular portion of the field and  $q$  is the probability of equal chance for all portions of the field, then the proper allowance for the inequality is the weighted average of plot yields in which the plot yield has a weight  $q/p$ . The application of this correction factor to the plot yields in the present scheme shows that the estimated average yield remains unaltered, the difference between the original estimate and the corrected one being less than half of 1 per cent.

11. A detailed set of instructions for the conduct of the proposed field work, namely (a) selection of fields, (b) location of plots, (c) harvesting and the connected operations, and (d) driage was prepared for the guidance of the staff. These will be given in the fuller report. Four forms of returns were prescribed under the instructions. In form 1 of returns, the Agricultural Assistant with the assistance of the Patwari concerned is required to show the Khasra numbers of all the wheat-growing fields in the selected village, arrange these serially and to select three fields out of these by using the random numbers supplied to him. In form 2 of the returns he is required to give, in respect of the three selected fields, the pairs of random numbers selected for locating plots, the type and level of the soil, preceding crop, preceding manure, manuring in current season, whether the field was irrigated or rain-fed, the variety of wheat grown, name of the crop mixed with wheat if any, proportion of wheat in mixture, estimated yield of wheat per acre and finally the date fixed for harvesting. The Agricultural Assistant was also asked to give detailed remarks in regard to the general condition of the crop. Both the forms 1 and 2 were required to be sent to me immediately after the work of selection was over. Under form 3 of the returns the Agricultural Assistant is required to fill in the actual results of harvesting in respect of each plot in a village. The Agricultural Assistant was specially asked to note any changes in the condition of the crop that might have taken place since the fields were selected. The return was required to be sent immediately after the harvesting in each village was over. Form 4 sets out the results of driage.

12. In a scheme of this nature the entire success of the survey depends on the reliability of field work, and, therefore, on the thorough

training of the staff, and consequently the greatest attention was paid to this aspect of the work in this scheme. The training of the senior staff was commenced on the 12th March at Lahore. All the Deputy Directors of Agriculture and the Extra Assistant Directors of Agriculture as also the Provincial Statistician and the Statistician attached to the scheme were present. The plan of work, the duties and responsibilities of each member of the field staff, the meaning and implications of the instructions drawn for the actual conduct of the field work, the manner in which the returns were to be filled and despatched, were all explained in detail to the staff. They were also told how to draw up the programme of work of the Agricultural Assistants under them and how to supervise and check the conduct of their work. The training was completed with a practical demonstration of the actual conduct of field work arranged in villages around Lahore. Before they dispersed, each of the Deputy Directors of Agriculture and the Extra Assistant Directors of Agriculture was supplied with the requisite number of copies of instructions for field work, copies of the various returns and the sets of random numbers and all the equipment required for carrying out the field work.

13. Immediately on arrival at their respective headquarters, the Deputy Directors of Agriculture in turn commenced the training of the Agricultural Assistants under them. In addition to the practical demonstration of field work in neighbouring villages, the junior staff were also assigned independent villages in the neighbourhood to test whether they had understood the work. In addition to these arrangements, the Statistician attached to the scheme and the Provincial Statistician of the Department of Agriculture went round from district to district training the field staff in the conduct of field work. Altogether it took about a month to complete the training and arrangements of work.

14. After receiving training the Agricultural Assistants went back to their respective tehsils and proceeded to locate villages selected for crop-cutting work. They did not have much difficulty in locating villages since they had already worked in the tehsils and had with them maps showing the detailed lay-out of their tehsils. After drawing up the programme of their work in consultation with the Extra Assistant Directors of Agriculture, the Agricultural Assistants visited the villages, one by one, selected three fields in each village as laid down in the instructions, and in consultation with the owners of the fields fixed suitable dates for harvesting. On the days appointed they went to the village concerned with their mukkadams and got the plots harvested, threshed, winnowed and weighed in their presence.

15. The list of fields growing wheat, pure or mixed with other crops, was copied from the records of the patwari. In nine out of the selected villages the *girdawari* was incomplete and so no experiment could be conducted in them. The fields were selected with the help of random numbers supplied. The plots were marked with the help of another set of random numbers supplied for the purpose. The whole procedure of selecting the fields and plots by means of random numbers was so arranged that it could be checked by examining returns at the centre.

16. Excepting in rare cases, cultivators of the selected fields readily agreed to have the experiments done in their fields and gave their whole-hearted co-operation. In most cases they arranged for the labour required for harvesting, and in return were paid a fixed amount towards labour charges. The Deputy Directors of Agriculture and the Extra Assistant Directors of Agriculture supervised the work and satisfied themselves that the fields and the plots were selected in accordance with the procedure laid down. As a rule, wherever possible, the operations of harvesting, threshing, winnowing and weighing were done on the same day, but in cases where the produce was moist, it was allowed to dry up under the care of the Assistant and the mukkadam and was threshed after a day or two in their presence. The produce obtained was invariably weighed with the help of weights which were standardised.

17. The returns filled in by the Agricultural Assistants began to reach me from the middle of April and continued to come in till about the beginning of June. The returns which were over 3,000 in number, were scrutinised in the office, as they came in. Our scrutiny showed that the information supplied, including that on the general condition of the crop at the time of selection and on the changes in the condition of the crop since the date of selection was detailed and complete. The information provided a good means of detecting dishonest work and gave a ready explanation wherever eye-estimates and the actual yields differed a good deal. Thus rains and hailstorms had damaged the crop during the interval between the date of selection and of harvesting; occurrence of disease like rust and blight during the interval was also reported from some places. In a few cases the grain was found to have developed during the interval resulting in larger yields than those expected at the time of selection. The position of the plot itself gave an explanation for the difference in the estimated average yield for the field and the actual harvested yield. Thus plots often lay in a portion which was either better or poorer in comparison with the average; often there were *kalar* patches in the plots; often they lay by the

roadside, were covered by shade of trees and so on. Often there was a path, a nala or a hill torrent going across the plot resulting in smaller yield per acre than would be available had the whole of the area been under wheat.

18. Out of a total of 2,208 experiments conducted under the survey, a small number had to be rejected on account of incomplete information. Thus from one village it was reported that adequate labour being not available, threshing and weighing could not be completed by evening. In another distant village it was dangerous to continue threshing and winnowing after 6 o'clock for fear of thieves and animals from the adjoining forest, and, therefore, the operations had to be left incomplete. In a third village the produce was tied up and left over, but was found to be stolen the next day. In a few cases the plots lay in fields which were subsequently described as *kharaba*, and were specifically reported to be excluded from the area under the wheat crop. In other cases, the crop, being green, could not be threshed on the same day; it was left over for a day or two to dry up, but the Agricultural Assistant having found it impossible to supervise the operations on the succeeding days, the returns were excluded from the analysis. In a few cases, the cultivators harvested the selected fields before the dates fixed for harvesting. In one village standard weights were not used for weighing the produce. Altogether 59 experiments were rejected on account of incomplete information. Data for the remaining experiments have formed the basis of our estimates and analysis.

19. The survey which is admittedly in the nature of experiment in method was conducted with a fourfold object; (*a*) to determine the average yield of wheat per acre and the total out-turn for the province with their sampling errors; (*b*) to collect ancillary information in respect of the acreage under different factors such as varieties, irrigation, manures, soil, previous crops and others; (*c*) to study the variability with a view to test the efficiency of the sample technique employed in the survey and to suggest improvements for reducing the sampling error in future years; and (*d*) to test the possibility of an alternative approach of eye-estimation on random fields for estimating the yield.

20. As all the data were received only a month ago, it has not been possible to undertake the study of items (*b*), (*c*) and (*d*). All that has been possible to do is to work out the average yield per acre, the total out-turn and their sampling errors. These are described in the following sections.

21. Table I shows the classification of selected fields as irrigated and rainfed, and as pure and mixed.



TABLE I  
*Number of Plots*

	Irrigated	Rainfed	Total
Pure ..	998	591	1589
Mixed ..	245	315	560
Total	1243	906	2149

It will be seen from the table that 58 per cent. of the selected fields are irrigated and the remainder rainfed. These percentages are in good agreement with the official figures for the percentage of area under irrigated and rainfed wheat, viz., 57:43, and demonstrate the representative character of sampling. The table also shows that roughly 20 per cent. of the fields under irrigated wheat and 35 per cent. of the fields under rainfed wheat are sown mixed with other crops mostly gram. The corresponding official figures for acreage are, however, not available, as no separate record is maintained of the area under pure wheat and of the area sown with wheat mixed with other crops. What is available is the total of the acreage under pure wheat and the area under wheat as separated by the village accountant from the mixed fields. In the official procedure of calculating the out-turn the whole of this area is considered as area under pure wheat and is multiplied by the figure for the average yield per acre as found from crop-cutting experiments on only pure wheat-growing fields. The calculation involves an assumption that wheat gives the same out-turn per acre whether it is grown pure or mixed with other crops.

22. In Table II are set out the calculated averages for the yield of wheat in maunds per acre from the returns of actual yields of wheat reported to us.

TABLE II  
*Average yield of Wheat in Maunds per Acre*

	Irrigated	Rainfed
Pure ..	12.3238	7.4886
Mixed ..	8.7508	5.6420

The first row of the table gives the average yield of wheat in maunds per acre for fields sown solely with wheat, while the second row gives the average yield of wheat only in maunds per acre from fields sown with wheat in mixture with other crops. The latter are naturally less than the former,

as wheat occupied only a portion of the area in mixed fields. The appropriate method of estimating the total out-turn is to multiply these several averages by the corresponding areas. These are, however, not available, and we have to look for other methods of utilising the data for estimating the total out-turn.

23. One alternative is to multiply the official figure for the area under wheat by the average yield of wheat calculated from experiments on pure wheat fields only. This method of calculating the out-turn is, in fact, the official procedure, but as has been pointed out above, it involves the assumption that wheat gives the same out-turn per acre whether grown pure or mixed. Moreover, from the statistical point of view we cannot justify the throwing away of this relevant information in respect of experiments on mixed fields when we have already spent time, labour and money on obtaining it. On the other hand, the limitations of official acreage statistics are a handicap in the way of utilising this information. Having foreseen this difficulty we collected information on the estimates of the proportion of area under wheat in mixed fields selected for the survey. This information is made use of in estimating the average yield of wheat per acre by dividing the actual harvested yield of wheat by the proportion of the area under wheat. Table II modified in this light is set out as Table III.

TABLE III

*Average Yield of Wheat in Maunds per Acre*

(The appropriate standard errors are given in brackets just below the corresponding averages)

	Irrigated	Rainfed	Total
Pure ..	12.3238 (0.1858)	7.4886 (0.2007)	10.5254 (0.1385)
Mixed ..	12.2264 (0.3827)	8.5576 (0.3099)	10.1627 (0.2417)
Total ..	12.3046 (0.1672)	7.8603 (0.1696)	10.4309 (0.1202)

The marginal figures represent the average yield in maunds per acre for different heads. The marginal averages under the irrigated and the rainfed classification multiplied by the corresponding areas give the out-turn of wheat for all the 27 districts covered by the survey. These are shown in Table IV.

TABLE IV

	Irrigated	Rainfed	Total
Average yeild in maunds per acre ..	12.3046	7.8603	..
Acreage under wheat ..	5,218,000	4,019,100	9,237,100
Out-turn in tons ..	2,375,550	1,145,930	3,521,000
Official estimate in tons ..	2,319,000	961,000	3,280,000

In the last row of the table is shown the official estimate for the out-turn of wheat in tons for the 27 districts. It will be seen that the sample estimate of out-turn of irrigated wheat is in close agreement with the official figures but that of unirrigated wheat is higher by over 19 per cent. than the corresponding official figure. The sample estimate of the total out-turn of wheat for the 27 districts is higher by 241,000 tons or by 7.4 per cent.

24. Whereas the estimate we have derived relates to the yield threshed the very day or the next on which it is harvested, the general practice in the province is to allow the harvested produce to dry in the field and/or on the threshing floor for about a week or two before it is actually threshed. The total out-turn is considered as that out-turn which is obtained from threshing the harvested crop in accordance with this general practice. We have, therefore, to allow for driage from our estimate. The results of reweighment of 5 seers of sample after a fortnight reported to us give an average driage of 2.0 chhatanks equivalent to 2.5 per cent. This we have to deduct to bring our estimate in line with the official estimate.

25. As against this some unavoidable losses have occurred during the process of harvesting, threshing and the connected operations. As the threshing had to be done on the day of harvesting to ensure the reliability of results, all the grain could only with great difficulty be separated from the ear-heads. Although every effort was made to separate all the grain, yet a little grain, particularly where the crop was moist, escaped in the ear-heads. The grain so left was excluded from the weighment. There also appears to be a long-established practice of leaving a few ear-heads in the fields for the menials of the village. Despite our instruction that every ear-head should be collected from the experimental plots, some losses have occurred on this score here and there. There is no means of estimating the loss on these heads, but we believe it was negligible. Assuming it to be one half of 1 per cent. we shall deduct 2 per cent. from our estimated out-turn and obtain the net estimate of 2,328,000 tons of irrigated wheat and 1,123,000 tons of unirrigated wheat as against the official figures of 2,319,000 and

is obtained in a round-about way by multiplying the so-called 'normal' with the condition factor of that year. We are reporting elsewhere the comparison of the results of the method of anna estimate with the actuals. Here it will suffice to state that 'anna' system is understood differently by different men, and leads to errors of very large magnitude. We have not had time to analyse the data for reporting on the personal bias of different workers in forming the eye estimates nor we have had time to work out the correlation between the eye estimate and the actual out-turn. The results reported above, however, show the possibility that if eye estimates by the agricultural staff are obtained on a large number of fields selected at random, it should be possible to estimate the actual out-turn with considerable precision.

28. The net increase of 5.2 per cent. revealed by the survey is small, but the increase of 17 per cent. in the out-turn of unirrigated wheat is rather large. In anticipating difference between the sample and the official estimates it has to be remembered that an experienced administrator can derive a fair idea of the average yield per acre for his district by conducting a few crop-cutting experiments even in accordance with the existing defective procedure. In the very nature of things, therefore, it is unlikely that his estimate can be wide of the mark by more than 15 to 20 per cent. margin of error. In a vast country like India, moreover, even a small percentage of difference means hundreds of thousands tons of production. It is, therefore, necessary to estimate the actual production with the maximum precision. The real advantage of conducting crop-cutting surveys in accordance with the procedure outlined in this report is that it achieves this object. There is another aspect which brings out the utility of the crop-cutting survey in the estimation of out-turn. There is a tendency in official estimates to keep close to the average of preceding years. The result is that the true fluctuations in the out-turn from year to year are not brought out. These fluctuations may be important in determining the policy and can be accurately assessed only by means of conducting crop-cutting surveys. The results of future surveys alone will show the advantage of crop-cutting surveys. In the meantime it is necessary to repeat the survey for perfecting the technique and for examining the possibility of an alternative and cheaper approach of forecasting yields. Such a survey incidentally will give us a picture of the agricultural conditions prevailing in different tracts of the country and over a period of years reliable estimates of normal yields for each district.

29. This is the first large-scale survey of its type carried out over an area of 88,000 square miles. The success of it is due to the whole-hearted co-operation of the Department of Agriculture in the Punjab. They can well be proud of their work, particularly as they will be giving a lead to other provinces in maintaining accurate statistics for their out-turns. The encouragement and the keen interest of the senior officials of the Government of India at the various stages of this survey contributed in no small measure to its success. Particular mention must be made of the help received from Sir P. M. Kharegat, Additional Secretary, Education, Health and Lands Department, Mr. B. Sahay, Deputy Secretary, Education, Health and Lands Department, Mr. D. R. Sethi, Director of Agricultural Production, Mr. H. R. Stewart, Vice-Chairman, Imperial Council of Agricultural Research, and Mr. S. M. Srivastava, Secretary, Imperial Council of Agricultural Research. It is hoped that the survey will be conducted year after year as recommended by the Inter-Departmental Committee on Agricultural Statistics, and would be carried out in the same spirit and with even greater efficiency than has been done during this year. The importance of the results of such surveys far outweighs the expense involved, which in itself happens to be only a little more than a thousand of rupees per district.