

## A SURVEY OF MAINCROP POTATOES

## I. ESTIMATES OF YIELD, 1948-50

BY G. V. DYKE AND P. R. D. AVIS  
*Rothamsted Experimental Station, Harpenden, Herts*

(With One Text-figure)

## 1. INTRODUCTION

The potato crop of Great Britain is the only major item of the national diet (except milk) produced almost entirely at home, and as storage from year to year is impracticable, it is important that the total production in any one season should be adequate to meet the demand until the next harvest. The cost per acre of potato growing is very high, in purchased material such as seed and fertilizers, and in labour and use of tractors and equipment. Further, a large proportion of the crop is grown on the better arable soils, on land of high rental values. The gross receipts per acre from potatoes are at least double those from cereals and are exceeded only on market-garden holdings. The cash value of the total crop as harvested is greater than that of any other single crop, although the acreage devoted to potatoes is comparatively small (in England and Wales wheat, oats and barley occupy each about twice the total potato acreage).

The maintenance of a high and steady level of yield is therefore of great importance, equally to the individual grower and to the consumer. Comparison of yields measured in field experiments and obtained on first-class commercial farms with the official estimates of yield for the country as a whole has for some time revealed a considerable discrepancy. It was therefore decided by the Agricultural Improvement Council that a survey should be carried out on maincrop potatoes with two objects:

(a) To test the official yield estimates against estimates derived from sample weights.

(b) To examine the practices of growers and in the light of results of experiments, to determine any improvements which might be practicable.

It was also hoped that the results of the survey might suggest useful new lines of experimental research.

In 1949 and later years, a scheme for obtaining August forecasts of yield from sample weights was tried out. It is hoped in later papers to give an account of this investigation and of the results of the survey concerning the methods used by potato growers. The present paper deals only with the problem of estimating the mean yield of an area such as England and Wales.

The survey was planned and carried out by officers of the National Agricultural Advisory Service in collaboration with the Department of Statistics, Rothamsted Experimental Station. The field records were completed by the county staffs, under the supervision of Provincial Crop Husbandry Officers, and sent to Rothamsted for analysis. The field work extended over the three seasons 1948-50, covering about 3500 fields in forty counties in all. The results of the first two seasons have already been summarized (Boyd & Dyke, 1950), and various duplicated reports have also been circulated.

## 2. SURVEY METHODS

Farms were selected as follows: a sample of farms was chosen in each participating county, the 'frame' used being a register of the acreages of potatoes which the farmers expressed their intention of growing in the current season. (These are given in returns made in March.) The farms were divided into three size groups on the basis of these acreages, and a variable sampling fraction in the ratio of 1 : 4 : 8 was used on the strata so formed. In some counties (in order to equalize the amount of work required of district officers) the procedure was applied to each administrative district separately.

Table 1. *Distribution of sampled fields*

Province	Total	Participating		
		1948	1949	1950
Northern	4	4 (280)	4 (126)	4 (160)
Yorkshire and Lancashire	4	4 (76)	4 (71)	4 (73)
East Midlands	7	7 (177)	7 (111)	7 (125)
West Midlands	6	3 (206)	3 (70)	4 (129)
Wales	13	2 (214)	3 (169)	3 (142)
South-western	6	6 (195)	5 (106)	6 (82)
South-eastern	10	6 (218)	7 (118)	6 (128)
Eastern	10	8 (206)	7 (162)	7 (214)
	60	40 (1572)	40 (933)	41 (1053)

The distribution of sampled fields among the advisory provinces and the numbers of counties participating is shown in Table 1.

Two visits were made to each selected farm, the first in June or July and the second in September or October (i.e. as late as conveniently possible before lifting). At the first visit the fields of potatoes actually planted were recorded, together with their acreages. If either one or two fields were being grown on a farm, further particulars, including variety, date of planting, source of seed, fertilizers, cultivation details, etc., were obtained for each field. If there were more than two fields, the information was obtained for a random sample of two fields. At the second visit four systematically located sample lengths of row, approximately 6 ft. long, were dug from each of the selected fields and the produce weighed (ware, seed plus chits, and haulm, separately). The exact length of row lifted was determined by the row width (measured at each sampling point) so that the area dug was constant. This was arranged so that a simple calculation using only the divisor 3 converted the total sample weight in pounds to the estimate of yield in tons per acre. A third visit was made after lifting to a subsample of one in ten farms, and sample areas were marked out on the selected fields and carefully dug by hand to ascertain the amount of potatoes remaining in the ground. On this same subsample of farms measurements were also taken at the first visit to determine the actual area of land under potatoes in each selected field so that an adjustment could be calculated to allow for uncropped areas (such as headlands) which are normally included in the official acreage.

### 3. ESTIMATES OF YIELD

#### (a) *The survey estimates compared with farmers' weighed yield*

The survey estimates of yield are based on weighed samples chosen in such a way as to relieve the officers concerned of the responsibility of making 'representative' choices, either of farms, of fields or of samples within fields. The magnitude of errors due to sampling at each stage may be estimated from the results themselves. The existence of bias (arising from bad sampling or causes outside the control of the samplers), can be tested by comparing yields calculated from the farmers' weighings of produce (where the crops were weighed shortly after lifting) and the sample yields of corresponding fields.

It is necessary to make various adjustments to the sample yields before comparing them with farmers' yields, as the latter are reckoned rather differently. First, the samples were carefully lifted by hand and few, if any, potatoes of ware size were left in the ground. The weights of tubers found in the sample areas forked over after potato harvest on the subsample show that substantial quantities of ware are wasted by inefficient lifting. On the

average 0.7–0.8 ton/acre of ware, together with 0.2–0.3 ton of smaller tubers, were recorded each year. It is possible that in a few cases subsequent harrowing and picking would somewhat reduce the final loss; in some cases a second picking was done because of the results of the sampling. The quantity of ware left in the ground represents nearly one-tenth of the crop grown.\*

Secondly, the grower normally calculates yield per acre from the total acreage of the field including headlands, etc., associated with the crop (the official estimates are based on this acreage); the survey estimates are of yield per cropped acre. The measurements on the subsample showed a mean difference of about 4% between the gross acreage and the area planted, leading to adjustments of 0.3–0.4 ton/acre. Thirdly, in 1950 only, an estimate was obtained of the quantity of ware at sampling which would be considered waste due to disease or greening.

In 1948 and 1949 the farmers' yields were obtained without reference to the grading of the potatoes: it was assumed that the official minimum ware riddle would be used by all growers, yields being based on potatoes standing on this riddle. In 1950 the riddle used was stated, and it was found that there was considerable variation in the riddle sizes used, though the majority of riddles used were  $1\frac{1}{4}$  or  $1\frac{1}{2}$  in. In 1948 the official minimum riddle was fixed at  $1\frac{1}{2}$  in. and to enable comparisons to be made with the sample yields, which were on a  $1\frac{1}{4}$  in. riddle (the official minimum in recent years except 1948), a series of trials was made mainly on survey fields to determine the mean weight of crop falling between the  $1\frac{1}{2}$  and  $1\frac{1}{4}$  in. riddles. This was found to be 0.3 ton/acre. This figure was verified from a smaller number of trials in 1949. Adjustments were made as follows:

1948. On all fields: 0.3 ton/acre and an additional (arbitrary) adjustment of 0.1 ton/acre (for the unknown proportion of cases where meshes larger than  $1\frac{1}{2}$  in. were used). *Total: 0.4 ton/acre.*

1949. For the unknown proportion of cases where riddles larger than  $1\frac{1}{4}$  in. were used an arbitrary adjustment of 0.2 ton/acre.

\* It should be noted that the total of ware and smaller tubers left in the ground is about equal in weight to the average rate of planting of potatoes. This confirms the observation (Doncaster & Gregory, 1948) that the density of volunteer plants in crops following potatoes is often high and occasionally comparable to the density of a planted potato crop. The effect of volunteers as potential sources of virus infection is dealt with by Doncaster & Gregory: other consequences such as the reduction of subsequent crops by competition from volunteer potato plants, or their maintenance of potato root eelworm infestations are not known with accuracy.

1950. For each field where a  $1\frac{1}{2}$  in. riddle was stated to be used the adjustment of 0.3 ton/acre was made. Fields where a  $1\frac{1}{4}$  in. riddle was used were accepted without adjustment and all others were rejected for the purposes of comparison with farmers' weighed yields. The mean adjustment is 0.2 ton/acre.

The agreement between the sample yields and the weighed yields in 1948 and 1949 is fair (Table 2), but it is evident that in 1950 a real discrepancy occurred, although the accuracy of the comparison is reduced because of the small number of comparisons available. When the individual discrepancies were examined, it was seen that there was considerable variation between counties and between districts within counties. This suggests that the procedure varied from district to district in such a way as to affect either the sample yields or the farmers' weighed yields. One possibility is the varying losses involved in periods of storage. Although the results from one county were omitted because many of the

(b) *Sampling errors of estimates of yield*

The mean yields calculated from the sample weights are subject to variation of two types. First, the yield per acre of a particular field, calculated from the total weight of the four sample lengths, has a standard error of about 12%, as estimated from the differences between the four individual weights. Secondly, the mean yield of an area (e.g. a county) as calculated from the sample weights of the sampled fields is subject to additional variation arising from the differences in yield between the potato fields of the county.\* (If samples were dug in *all* potato fields this variation would be eliminated). The standard error due to both causes combined, expressed in units of a single field, is approximately 40% of the yield. Thus if the unweighted mean yield of the sampled fields in a county is 10 tons acre the standard error of the sample estimate is  $4/\sqrt{n}$  tons/acre, where  $n$  is the number of sampled fields.

Table 2. *Comparison of farmers' weighed yields and sample yields, all in terms of  $1\frac{1}{2}$  in. ware (tons/acre)*

	1948	1949	1950
Number of fields	136	72	46
Mean yield (gross) from samples	11.3	8.9	11.4
Deduct: Ware left in ground	0.8	0.7	0.8
Correction for acreage	0.4	0.4	0.3
Ware unsaleable	0.2*	0.2*	0.3
Mean yield (net) from samples	9.9	7.6	10.0
Mean of farmers' yields	9.5	7.8	8.9
Add: Correction to $1\frac{1}{4}$ in. riddle	0.4	0.2	0.2
Total	9.9	8.0	9.1
Mean excess of sample yields over weighed yields	0.0	-0.4	+0.9
Standard error of mean excess	$\pm 0.2$	$\pm 0.2$	$\pm 0.3$

\* Not obtained from samples; the estimates given are based on the fact that blight was more serious in 1950 than in 1948 or 1949.

growers' yields were based on weighings in late November or December, it is likely that a few such cases have escaped notice. The infection of blight which was well established in the tubers of many crops at the time of sampling in 1950 is likely to have increased rapidly and contributed to this loss; an example in the Eastern Province was a report of 5 tons/acre of blighted tubers.

It seems, therefore, that if allowance is made for discrepancies due to various causes, the sampling method used is capable of giving substantially unbiased estimates of yield, though the losses after the sampling date may be difficult to deal with.

Losses during storage in clamps were not examined in the survey, although such losses do in some years seriously reduce the quantity of potatoes reaching the market in the latter part of the season. The main reasons for the losses are blight, rots and bad clamping, but no great amount of information is available on the subject.

When weights are introduced to allow for the different acreages of the sampled fields and for the different sampling fractions used both in selecting farms and fields, the standard errors of the weighted means are greater than those of the unweighted by a factor of about 1.1. This is the equivalent of a reduction of 20% in the number of fields sampled. Further, the standard error of the mean yield of all surveyed counties, allowing for the different potato acreages, is appreciably inflated by the distribution of samples between counties *not* in proportion to their potato acreages. The uneven distribution of samples between the sampled counties has the

\* It should be noted that the variation between fields on the same farm is appreciably smaller than between fields on different farms in the same size group and county. The precision obtained by sampling two fields per farm on a given number of farms in a given area could be equalled by sampling one field per farm on about one-fourth more farms.

effect of reducing the effective number of samples by about one-half to two-thirds. This loss of precision could have been entirely avoided by ensuring that each county undertook the number of samples appropriate to its potato acreage. A considerable loss of precision can arise in this way from a very small number of counties where the number of samples per thousand acres of potatoes is much lower than elsewhere. This emphasizes the need for co-ordination of the effort devoted to surveys whose objects are estimates of mean yield, etc., when several administrative areas are involved.

Table 3. *Standard errors of estimates of mean yield of all sampled counties (tons/acre)*

Year	Standard error per field (all sources)	No. of fields	Standard error of weighted grand mean
1948	4.0	1572	0.16
1949	3.3	933	0.17
1950	3.3	1053	0.16

Table 3 shows the standard errors of the means of all sampled counties in each year. The weighted means over all sampled counties have standard errors of about  $\pm 0.16-0.17$  ton/acre in each year.

(c) *The official estimates*

The official estimates of yield issued by the Ministry of Agriculture are based on reports by officers of the N.A.A.S. giving estimates for each parish (Britton & Hunt, 1951). In 1948, and earlier years, this work was done by 'crop reporters' specially appointed by the Ministry for the purpose of estimating yields. These parish figures are arrived at after consultation with 'representative farmers', and are intended to take into account the proportions of crops having different levels of yield. As far as is known, no guidance is given as to exactly how the information obtained from representative growers and the judgement of the officer himself are to be combined to give the parish mean. No estimate of the magnitude of errors of estimation, systematic or otherwise, is possible. The fact that *The Times* estimates, which are presumably compiled in a similar way, consistently tend to be higher illustrates the possibility of bias in such estimates.

(d) *Comparison between official and sample estimates*

Table 4 shows the weighted mean sample estimates for all sampled counties for each year of the survey, compared with the weighted means of the official estimates for the same counties. In order to present the comparison in terms of ware potatoes over a  $1\frac{1}{4}$  in. riddle the official estimates of total produce have been diminished by the mean sample estimates

of seed and chats. The official estimates are, for each year, well below the estimates from the survey. The mean discrepancy is 1.8 tons/acre or about one-sixth of the survey estimates.

It should be noted here that the official estimates refer to 'maincrop and second early' potatoes, while the survey was in theory confined to 'maincrop' only. In practice, however, the officers selecting farms for survey did not know what proportion (if any) of 'second earlies' was grown until the farms were visited. No case was recorded in 1948 or 1949 of either farm or field being rejected because of the growing of 'second earlies' only, while many of the sampled fields were recorded as carrying varieties generally classed as 'second early'

Table 4. *Comparison of sample estimates with official estimates (tons/acre)*

	1948	1949	1950
Official estimates (excluding first earlies; sampled counties only)	8.1	7.2	8.3
Deduction for seed and chats (estimated from the survey results)	0.3	0.4	0.5
Net yield of ware			
Official estimates	7.8	6.8	7.8
Sample estimates	9.5	8.5	9.9
Excess of sample estimates over official estimates	+1.7	+1.7	+2.1

(the only official distinction seems to be based on varieties). In 1950, fieldmen recorded for each farm the acreage of 'second earlies' (defined as potatoes intended in the early part of the season for lifting before 1 September). Although an average of 5% of acreage was recorded under this heading the fields were very seldom, if ever, rejected for that reason. It appears, therefore, that the two sets of estimates, official and survey, do in fact refer to almost exactly the same set of fields. The difference between yields of 'first early' and 'second early and maincrop' as given by the official estimates is about 3 tons/acre; it is likely that the mean difference in yield between 'second early' and maincrop is not more than 1 or at the most 2 tons/acre. The disturbance to the sample estimate from the exclusion of a small acreage of 'second early' potatoes is therefore unlikely to amount to 0.1 ton/acre.

The official and sample estimates for individual counties for each year are shown in Fig. 1 (the mean sample estimate of the yield of seed and chats in each year has been deducted from the official estimates). The full lines in the diagram represent the regressions for the 3 years separately of the official estimates on the sample estimates, adjustments being made for the sampling variance of the sample estimates. The broken line represents equality between estimates of the two types.

The three regression lines are remarkably consistent. (It must be remembered that the official estimates for 1948 were compiled by the crop reporters, whose duties passed in 1949 to officers of the N.A.A.S.). In counties where the sample estimates were high (10 tons/acre or above) the official estimates generally fall short of the sample estimates by 2 tons/acre or more. On the other hand, where yields were 5 to 6 tons/acre, agreement between the estimates was reasonably good.

Venn (1926) pointed out that the official estimates,

have been improved by a better distribution of samples among counties.

The results point to underestimation on the part of the official estimates, in each of the 3 years, especially in the case of high yields in particular counties, and in particular years. The discrepancy between the official and the survey yields is of the order of  $1\frac{1}{2}$  tons/acre, after all necessary corrections have been applied to the survey yields.

The experience gained in the survey indicates that the method of sampling adopted provides an

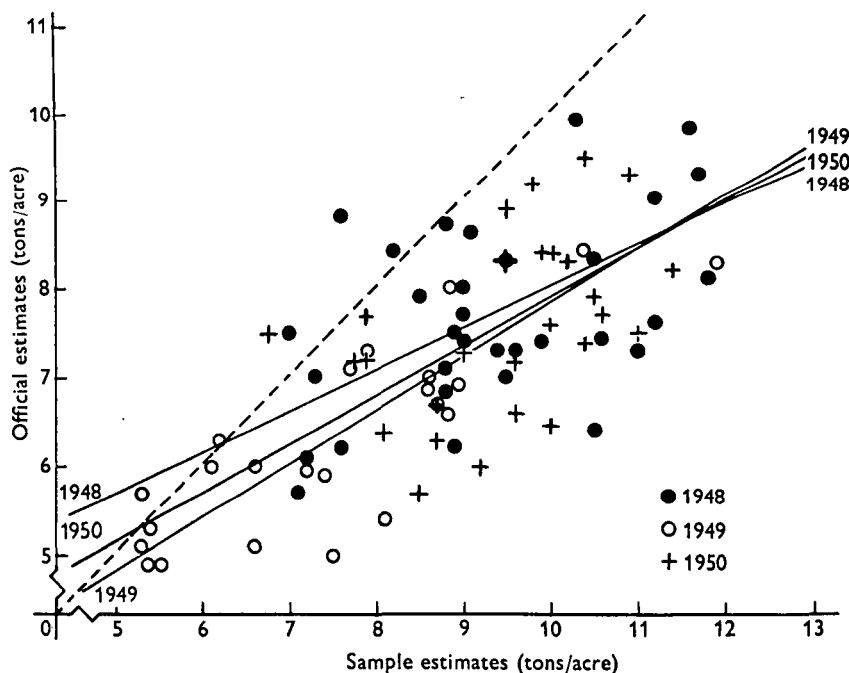


Fig. 1. The relation between the official and sample estimates of counties, 1948, 1949 and 1950. (Counties with less than 20 sampled fields excluded.  $1\frac{1}{2}$  in. ware.)

though comparatively accurate in years of generally low or average yield, regularly underestimate when the mean yield is above average. It seems likely from Fig. 1 that this underestimation arises from gross underestimation of counties with above average yields and not from a general underestimation of all yields.

#### SUMMARY

The survey shows that objective estimates of the yield of maincrop potatoes can be obtained from small samples carefully selected and dug by hand. Samples taken from about 1000 fields gave estimates of the mean yield of all counties sampled with a standard error due to sampling of less than  $\pm 0.2$  ton/acre. The precision of the estimate could

accurate and reliable method of estimating the yields of potatoes which could supplement, and, possibly, ultimately replace the present official estimates if more accurate estimates are required. A national scheme, properly designed, which would include all the potato-growing areas in due proportion should not be unduly expensive to operate. Estimates so obtained would not only be generally more accurate than those obtained by the present official method, but, perhaps more important, would indicate far more closely the fluctuation in yield from year to year.

Thanks are due to Dr J. H. Western, Plant Pathologist, Northern Province, to the following Crop Husbandry Officers: Mr C. D. Price, Northern

Province; Mr W. W. Gatenby, Yorks and Lancs Province; Mr S. J. Travers, East Midland Province; Dr D. H. Robinson, West Midland Province; Mr A. J. Davies, Wales; Mr R. Holliday, Eastern Province; Mr C. V. Dadd, Eastern Province; Mr J. N. Sharrock, South Eastern Province; Mr W. Q. Connold, South Western Province. Also to the District Advisory Officers concerned, in carrying out the detailed field work of the survey.

## REFERENCES

- BOYD, D. A. & DYKE, G. V. (1950). *N.A.A.S. Quart. Rev.* **10**, 47.  
BRITTON, D. K. & HUNT, K. E. (1951). *J. R. Statist. Soc.* **114**, 59.  
DONCASTER, J. P. & GREGORY, P. H. (1948). The spread of virus diseases in the potato crop. *A.R.C. Rep. Ser.* no. 7.  
VENN, J. A. (1926). *Econ. J.* **36**, 394.

(Received 16 January 1953)