Validation of Objective Method of Estimating Soybean Yield

By Charles E. Rogers and Douglas E. Murfield

TIMELY, RELIABLE measurement of annual crop production has assumed growing importance throughout the world in recent years. Because direct measurement is usually not possible, it is necessary to estimate acreage and yield, the components of production. The technique of objective yield measurement by harvesting small plots has been used since about 1940. The first extensive investigation of plot size and shape was conducted in India on wheat. Plots of about 12 square feet resulted in serious overestimation which decreased as plot size was increased. The overestimation was attributed to border bias, i.e., the tendency for samplers to include plants adjacent to but outside the plot. Measurements of other crops have shown similar bias.

In the United States considerations of variability of plant yields and costs have dictated two very small plots per sample field, and procedures to minimize border bias have been developed. The plot being used in soybean yield work consists of two 3-foot row sections, located in adjacent rows. By counting a random number of rows and paces in the sample field, the sampler reaches the plot location. When he has taken the prescribed number of places in the proper row, he places a thin dowel stick across two rows at the end of his toe. This marks the starting point for the 3-foot row sections. The length of the plot is measured in each row by sliding a steel frame into the plants in the row. This frame was designed to measure exactly the 3 feet of row to determine precisely the plants that lie within the row section. The second plot is similarly located. For the preharvest estimate of yield, the sample plots are harvested, the pods are threshed and weighed, the moisture content is measured, and the yield is computed. After harvest, other plots are laid out and gleaned to measure harvesting losses.

This validation study was conducted to determine whether or not there is appreciable bias in soybean objective yield procedures. Regular sampling personnel were employed and the same field procedures described above were used except for the number of sample plots per field.

Procedure

The Soybean Objective Yield Procedure Validation Study was conducted in the fall of 1964 in Illinois and Mississippi. For this study, 5 fields in Illinois and 7 in Mississippi were purposively selected on the basis of location, size, and expected yield. The selected fields, which ranged in size from 18 to 50 acres, were measured accurately with a surveyor's chain to obtain the acreage to be harvested. In each field, using the procedures prescribed for the regular soybean objective yield work, 100 sampling units (each a 2-row, 3-foot plot) were randomly selected and identified. Then the beans from these sampling units were harvested and sent to the regional laboratory where they were threshed and weighed, and their moisture content was determined. Within a few days after the farmer had harvested the field, post-harvest gleanings were made on 50 plots per field. Each of these plots was located 5 paces beyond an even- or odd-numbered preharvest plot and consisted of a 3-foot row section with its associated middle. Gleanings obtained from these plots were also mailed to the laboratory for threshing, weighing, and moisture testing.

As the fields were harvested by the farmer, each load was weighed and moisture tests were made. The yield for the field was computed by dividing the weight of beans harvested from the field by the measured acreage.

A sample estimate of gross yield was computed for each field from the threshed weight of beans (adjusted to harvest moisture content) harvested in the sample plots and the average

Reprinted from Agricultural Economics Research, Vol. XVII, No. 3, July 1965
by Standards and Research Division, Statistical Reporting Service.

row width. An estimate of harvesting loss was derived similarly from the postharvest gleanings data and this was deducted from gross yield to obtain an estimate of net yield. Within-field variances were computed for adjusted bean weights and gleanings from the sample plots, and were used to determine the variance of the net yield per acre. For this, the two variances were summed, since the covariance calculated for several fields was near zero and this indicated there was virtually no correlation between the gross yield and harvesting loss within fields.

Analysis

In this study, the null hypothesis being tested was that the yield estimated on a weight basis obtained by harvesting sample plots following regular field procedures does not differ from the yield on a weight basis as determined by harvesting the entire field and hauling the beans to the nearest scale for weighing. Since the experiment was designed to test field procedures for bias, a wide range of conditions was selected, and the averages and variances which were pooled for all fields are of primary interest. Table 1 summarizes the yields and standard errors computed for the different fields, as well as these values pooled by States.

Table 1.--Soybean yield and acreage data by fields and States, 1964

State and field	Measured acreage in field	Net yield		Difference	Standard error	Sampling error
		Harvested by farmer	Estimated from sample plots			
Illinois:	*Acres*	*Bu.*	*Bu.*	*Bu.*	*Bu.*	*Pct.*
1.....	24.6	32.91	31.90	-1.01	.800	2.51
2.....	22.6	20.92	21.19	+.27	.783	3.70
3.....	32.6	33.42	33.10	-.32	.967	2.92
4.....	18.9	25.85	28.53	+2.68	.479	1.68
5.....	27.9	18.40	18.84	+.44	.585	3.10
Total..	126.5	26.66	26.93	+.27	.357	1.33
Mississippi:						
1.....	50.4	29.54	28.08	-1.46	.901	3.21
2.....	29.6	22.67	23.71	+1.04	.932	3.93
3.....	27.2	41.07	40.89	-.18	1.445	3.53
4.....	22.8	9.65	10.37	+.72	.510	4.92
5.....	24.2	14.09	13.19	-.90	1.278	9.69
6.....	20.9	19.02	22.85	+3.83	1.247	5.43
7.....	19.4	19.65	20.10	+.45	.864	4.30
Total...	194.5	23.73	23.93	+.20	.412	1.72
Two States Combined..	321.0	24.874	25.096	+.222	.287	1.14

Using the t distribution to test the significance of differences between measured and estimated yields, we obtain a t value for Illinois of 0.76, for Mississippi of 0.49, and for the combined States of 0.77.

None of these t values are significant at the 67 or 95 percent level. The low magnitude of these t values indicates that any difference between the two yields could easily be explained by sampling error.

Confidence intervals may be set at the 95 percent level for each State and the combined States as follows:

State	Interval	Harvested yield	Estimated yield
Illinois.....	25.95-27.37	26.66	26.93
Mississippi.....	22.91-24.55	23.73	23.93
Two States combined.....	24.30-25.44	24.87	25.10

In all three instances, the estimated yield is well within the confidence interval derived from the estimated variance and the population harvested yield. Of the differences between the estimated and harvested yields by fields, five are negative and seven are positive. This suggests that no constant bias exists in one direction. The differences between yields were nearly offsetting, with a difference of only 0.27 bushel in Illinois and 0.20 bushel in Mississippi. The harvested yield for all fields except one in each State is within the 95 percent confidence interval. For these two fields, this suggests either (1) two unlikely combinations of sampling units occurred, or (2) there existed some departure in procedure which produced the larger differences. The fact that both these significant differences were positive suggests the possibility of a departure from procedure. However, if such differences occur rather infrequently, it may be that they are associated with some field or yield characteristic.

Conclusion

The small differences between estimated yields and harvested yields are not significant and may be attributed to sampling error. The hypothesis of no difference between estimated and measured yield is not refuted, and the conclusion follows that any bias in field procedures is insignificant and negligible.