# Pre-Harvest Estimate of Yield and Sugar Percentage Based on Random-Sampling Technique

H. E. BREWBAKER AND H. L. BUSH[1]

An accurate pre-harvest estimate is of considerable value from several viewpoints. First, the agricultural department likes to be in a position to give to the management an accurate estimate of the crop to be harvested. At the present time the department has been asked to make a crop prophecy from a personal observation of the fields. Aside from the pride of the agricultural department in being able to give a reasonably close estimate of the crop to be harvested, there are some practical points to be considered. In areas where the company has more than one factory and where they decide to equalize the length of the campaign between their factories to limit the losses from storage in piles, it is desirable that a reasonably accurate estimate be known, so that diversions can be made between those factories at the lowest net cost per ton for transportation. Inaccurate estimates may cause diversions to be made at greater cost than is finally justified when the true figures are known. The management also desires to know the length of campaign in order to provide operating supplies in sufficient quantity and to avoid over-purchases. Another very important point is that the estimated tonnage will give the management an estimate of sugar and by-product production from which output they will determine their sales policy. An accurate estimate of the tonnage to be harvested is also of value in determining the length of the campaign and from that to determine the date for beginning operations.

Since it is to be a pre-harvest estimate it really consists of two estimates: First, that of the yield and percentage of sugar at the time the samples are taken ; and, second, a prediction as to what will happen to these characters before harvest is completed. The latter must be, in some measure, at least, a matter of personal or group judgment, based on an intimate knowledge of the condition of the crop. However, as pointed out later in this paper, two dates of sampling provide some real measure of the trend. The magnitude of certain losses at harvest discussed more fully in another section of this report, will become recognized after the method has been used a few seasons and may be taken into consideration in predicting final figures for the completed harvest.

[1]Agronomist and Statistician, respectively, Experiment Station, Great Western Sugar Company.

In devising a scheme for arriving at a pre-harvest estimate we have considered three questions.[2] We will attempt to answer only the first two; the last question being left for those company officials whose duty it is to provide beets for processing.

1. Is the method legitimate or sound?
2. Will it be precise and accurate?
3. Will the cost be excessive?

A randomized-sampling scheme, including geographic stratification, was conducted by the Agricultural Marketing Service of the United States Department of Agriculture cooperating with the Iowa State College, in making pre-harvest estimates of wheat yields. King and Jebe (1) in a preliminary report on this scheme concluded, from data taken in North Dakota in 1938, that route sampling of the wheat crop does give a practical and efficient estimate of the yield per acre.[3] They found that stratification by varieties would have given a marked gain in accuracy, while a geographical stratification would have added little to the information. King and McCarty (2) in a report on pre-harvest sampling of the wheat crop in Oklahoma, Kansas, Nebraska, South Dakota, and North Dakota for 1939-1940 suggest that double stratification by variety and district might be advantageous in some years.

It is pointed out in both of the above mentioned reports, that by increasing the number of samples per field a larger number would need to be taken to obtain the same accuracy than if the number of samples were increased by sampling more fields.

In an attempt to make a pre-harvest estimate of the sugar-beet yield and percentage of sugar in the Great Western area, similar basic principles were employed in setting up a sampling scheme as were used in the studies reported above on wheat-yield estimates.

## Methods Used in This Study

The factory district was chosen as the basic unit, for which considerable accuracy was desired. The sampling was organized and carried through in 16 factory districts, 12 of which were in Colorado, 3 in Nebraska, and 1 in Wyoming. There was no particular interest in obtaining an accurate estimate for individual farms. Yield, percentage of sugar, and stand were considered.

Farms were listed for each factory district in order of their 3-year average beet yield for 1938-1940 with the current year's acreage in beets attached. All farms with less than 5 acres in beets for the current year were eliminated, since such farms constitute a very small percentage of the acreage in the Great Western territory.

[2]The proposed scheme was discussed with G. W. Snedecor, Director of the Statistical Laboratory at Iowa State College, to whom we are indebted for valuable criticism relative to the plan and analysis of the data.

[3]Figures in parentheses refer to Literature Cited,

The sampling problem was much simplified by leaving them out of consideration. The farms having 5 acres or more for the current year and immediate 3-year previous beet records were then listed or "stratified" into 5 classes by starting at the top of the list and arbitrarily making sub-divisions so that 20 percent of the acreage fell in each class. Approximately an equal number of farms for sampling purposes were then chosen at random from each class. The number of farms thus selected from each class was in accordance with the approximate ratio of 1 farm for each 100 acres. In eases where there was more than one field on the farm, the larger field was sampled.

Sampling of the randomly selected fields was then made according to a preconceived scheme. A sample consisted of 10 feet of row with 2 samples being taken from each field, regardless of size of field, on each of 2 sampling dates. The 2 samples for each date were taken in order that some information relative to the sampling error might be determined. The 2 samples were located as follows: One each in rows 40 and 80 from the nearest corner of the field, the sample to start at 100 paces from the end of the field, and to end 10 feet down the row. When the prescribed location for either sample fell out of the field, a substitute location within the field was to be determined on some other random sample basis. No choice was permitted in locating these points of sampling. If either end of the 10-foot length of row fell on the center of a beet, a coin was flipped to determine if it were to be included or excluded. All variations in stand, regardless of cause, were accepted as a part of the fundamental concept of random sampling.

Samples for the second sampling date were taken by going to the same locations as had been chosen for the first date, then skipping 10 feet in the same row as previously sampled, and measuring off the next 10 feet of row for this sample.

The first date for sampling was from September 2 to 8 and the second date, September 22 to 29. The estimated yields for these 2 dates were plotted and the curve drawn through these 2 points from 0 at the planting date and projected into October to give some basis for predicting the additional growth after the last sampling date. The error due to weak stands at the end of the row was not estimated on this basis, but is included with other practical discrepancies such as loss of beets covered with tops thrown in the field by the labor. There were also these other losses to be considered: Loss of beets in transportation to the dump; loss due to exposure after pulling, and piling and loss due to the labor topping some beets too low, since the beets taken for the samples were topped according to company tare standards. The total extent of these losses will be discussed later in this paper. It is to be assumed that differences in actual mean yields of the fields sampled from the sample estimate will be due to considerable extent, at least, to the above mentioned sources of error.

The factory stand figures at harvest were obtained in the usual way, through tare samples which are taken at frequent intervals for each contract. The beets were counted, weighed, and later calculated into a percentage of stand, using average weight per beet from the tare samples, and final tonnage delivered for the contract.

## Results

For the 16 factory districts, 1,317 fields were sampled with 2 samples each, or a total of 2,634 samples for each date. The mean percentage of stand, tons per acre, and percentage of sugar for each date of sampling, and the actual harvest results are presented by factory districts in table 1. The mean pre-harvest figures for the 12 Colorado factories, 3 Nebraska factories, and 1 Wyoming factory are plotted in figures 1, 2, and 3, respectively.

Perhaps the most striking observation to be made from the data presented in table 1 is the remarkably high correlation between the actual harvest results for the farms sampled and the corresponding factory averages. The largest discrepancies between these figures

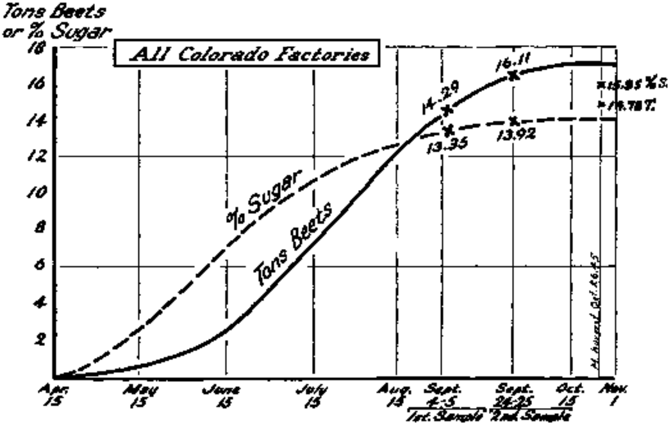


Figure 1.—Mean of pre-harvest samples for the 12 Colorado factories plotted in a growth curve, which is projected through the mean harvest date of October 20.45, and which is largely hypothetical except for the two actual pre-harvest points. The final harvest figures for these same farms are located on this graph for the mean harvest date.

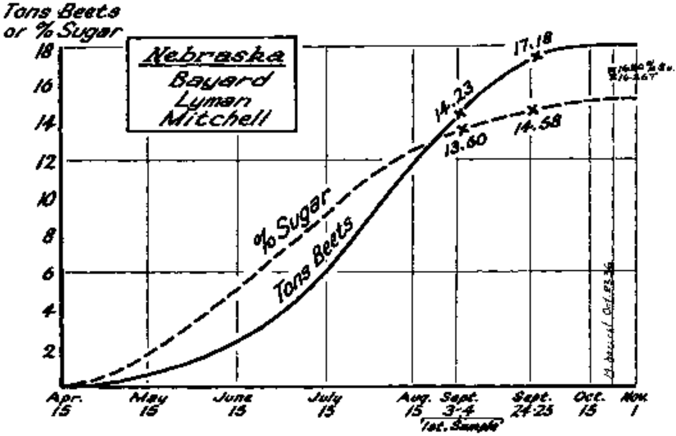


Figure 2.—Mean of pre-harvest samples for the 3 Nebraska factories plotted in a growth curve, which Is projected through the mean harvest date of October 23.30, and -which is largely hypothetical except for the two actual pre-harvest points. The final harvest figures for these same farms are located on this graph for the mean harvest date.

were: At Loveland, 5 percent in stand, 1.05 tons beets at Brush, and 0.23 percent sugar at both Lyman and Brighton. These factory variations tend to equalize out to the extent that for either of the Colorado or Nebraska averages, or for the average of all districts, the actual harvest figure for the farms sampled agrees accurately with the factory averages for all beets received. This high correlation is sufficient evidence that the farms chosen for pre-harvest estimates represented a nearly perfect sample of the Colorado or Nebraska districts or of the entire 1.6 factory districts including Lovell, Wyoming.

Since it has been shown that the farms chosen by random were a true sample, the difference between the last pre-harvest sample figures and final harvest results can be limited to three possibilities:

1. Errors in sampling the farms, which may in this case include a slight error resulting from failure to sample ends of fields.

2. Changes in yields or percentage of sugar between the last pre-harvest sample and the final completion of harvest.

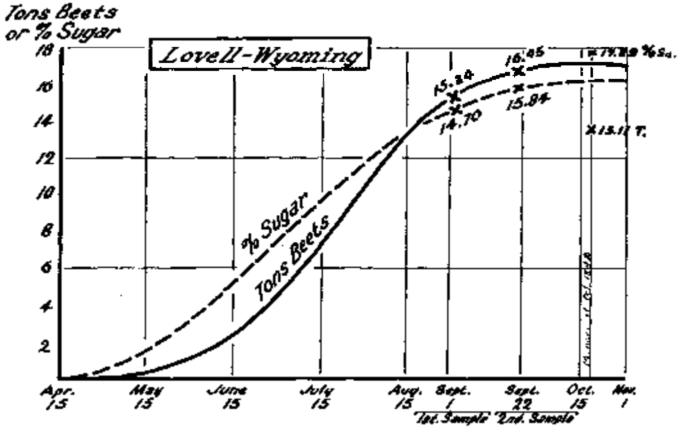


Figure 3.—Mean of pre-harvest samples for the Novell, Wyoming, factory plotted in a growth curve, which Is projected through the mean harvest date of October 18.42, and which is largely hypothetical except for the two actual pre-harvest points. The final harvest figures for these same farms are located on this graph for the mean harvest date.

3. Actual losses incident to harvest such as loss of whole beets In the field or along* the road, lower topping- by the labor than was used on pre-harvest samples, and from exposure during harvest.

As will be shown later (table 2), the variance "within fields" was comparatively small, and this error would certainly tend to average out for all fields sampled. Furthermore, the yields for each factory show fairly consistent gains in yield between the two sampling dates, and the curves (figures 1, 2, and 3) indicate consistent and reasonable improvement in each case. The percentage-of-sugar change during this period between the two pre-harvest sampling dates is probably not as good an index for this year as yield because this period was very rainy, and leaf spot was heavy in some areas. Both of these conditions are definitely unfavorable to sugar percentage increase. However, even under these conditions sugar percentage showed a definite, although not normal, improvement for all factories except Brush and Ovid, where there was heavy leafspot infection. It seems probable, for these reasons, that the averages of the

Table 1.—Sample estimates and final results for percentage stand, tons per acre, and percentage of sugar for all factory districts.

| | | Percentage Stand | | | | Tons per Acre | | | | Percentage of Sugar | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | For farms sampled | | | | For farms sampled | | | | For farms sampled | | | |
| Factory district | No. of farms sampled | 1st Samp. | 2nd Samp. | Harv. | Factory Average | 1st Samp. | 2nd Samp. | Harv. | Factory Average | 1st Samp. | 2nd Samp. | Harv. | Factory Average |
| Bayard, Nebraska | 85 | 82 | 82 | 70 | 72 | 13.00 | 16.10 | 14.83 | 14.70 | 13.20 | 14.13 | 16.07 | 16.00 |
| Lyman, Nebraska | 50 | 92 | 89 | 74 | 74 | 15.35 | 18.52 | 17.60 | 17.95 | 14.39 | 15.42 | 17.50 | 17.27 |
| Mitchell, Nebraska | 45 | 83 | 79 | 70 | 71 | 15.30 | 17.74 | 17.36 | 17.30 | 13.20 | 14.48 | 16.19 | 16.30 |
| Brighton, Colorado | 61 | 97 | 91 | 83 | 70 | 13.70 | 15.62 | 15.00 | 14.07 | 14.61 | 14.64 | 17.02 | 16.81 |
| Brush, Colorado | 62 | 86 | 84 | 70 | 73 | 15.58 | 17.42 | 14.54 | 15.59 | 12.80 | 12.80 | 15.80 | 15.70 |
| Eaton, Colorado | 120 | 94 | 92 | 77 | 79 | 16.13 | 17.28 | 16.02 | 16.43 | 13.24 | 13.85 | 15.96 | 16.01 |
| Ft. Collins, Colorado | 90 | 88 | 87 | 79 | 80 | 12.58 | 14.43 | 13.49 | 13.33 | 12.47 | 14.15 | 16.14 | 16.08 |
| Ft. Lupton, Colorado | 65 | 88 | 90 | 81 | 82 | 14.43 | 16.77 | 15.35 | 15.42 | 14.12 | 14.60 | 16.70 | 16.70 |
| Ft. Morgan, Colorado | 99 | 88 | 86 | 77 | 73 | 15.45 | 17.29 | 16.73 | 15.99 | 13.39 | 14.06 | 15.64 | 15.65 |
| Greeley, Colorado | 100 | 91 | 87 | 78 | 80 | 14.90 | 15.50 | 14.58 | 14.38 | 13.71 | 14.37 | 16.10 | 16.20 |
| Longmont, Colorado | 90 | 97 | 93 | 85 | 84 | 14.71 | 16.35 | 14.87 | 14.60 | 14.25 | 14.42 | 16.12 | 16.20 |
| Loveland, Colorado | 75 | 92 | 90 | 81 | 86 | 13.16 | 14.53 | 13.86 | 13.78 | 13.37 | 14.10 | 16.02 | 16.06 |
| Ovid, Colorado | 115 | 90 | 86 | 78 | 75 | 13.03 | 15.60 | 13.39 | 13.35 | 12.96 | 12.74 | 15.12 | 15.09 |
| Sterling, Colorado | 100 | 87 | 89 | 81 | 82 | 12.55 | 14.92 | 13.92 | 14.13 | 13.23 | 13.80 | 15.70 | 15.05 |
| Windsor, Colorado | 80 | 82 | 86 | 76 | 75 | 15.13 | 17.94 | 15.63 | 16.03 | 12.64 | 13.82 | 16.02 | 16.00 |
| Lovell, Wyoming | 80 | 100 | 97 | 74 | 71 | 15.24 | 16.46 | 13.11 | 13.13 | 14.70 | 15.84 | — (a | 17.29 (b |
| Average all districts (c | — | 91.5 | 88.3 | 77.0 | 77.5 | 14.24 | 16.28 | 14.88 | 14.89 | 13.47 | 14.13 | 16.03 | 16.02 |
| Colorado average (c | — | 90.1 | 88.6 | 78.3 | 78.9 | 14.29 | 16.11 | 14.78 | 14.79 | 13.35 | 13.92 | 15.95 | 15.95 |
| Nebraska average (c | — | 85.0 | 83.2 | 71.1 | 72.3 | 14.23 | 17.18 | 16.20 | 16.25 | 13.60 | 14.58 | 16.50 | 16.43 |

a) Individual tests not made.

b) Cossette average—not included in "Average all districts."

c) Weighted for number of farms in each factory district.

pre-harvest samples were quite accurate indices of the actual stand, yield, and percentage of sugar at the time the samples were taken.

Because of unusual and rather general rains during the latter part of September, a larger than normal increment in yield, and smaller than normal increment in percentage of sugar would have been expected between the two pre-sampling harvest dates. The curves (figures 1, *2,* and 3) were extended, assuming about average improvement following the last pre-harvest sample and during harvest. The final harvest averages for the farms sampled show lower yields for the three respective districts of Colorado (12 factories), Nebraska (3 factories), and Wyoming (1 factory) of 1.32, 0.9, and 3.32 tons per acre than was obtained for the second pre-harvest estimate. The hypothetical graphic projection of the yield curves through harvest would increase these apparent losses in yield to about 2.32, 1.74, and 3.99 tons per acre for these respective districts, this increment in yield being entirely reasonable and expected since growing conditions were relatively favorable well into the harvest period.

During the harvest period, conditions favorable to sugar storage improved with much reduced precipitation, and the increase in sugar percentage was probably better than average with apparent increases of 2.03, 1.85, and 1.45, respectively, for Colorado, Nebraska, and Wyoming districts, between the last sampling date and the final factory harvest.

In stand, there appears to have been a loss of about 2 to 3 percent between the first and second pre-harvest estimates, and from 10 to 12 percent between the second pre-harvest estimate and the final harvest figures. We have no explanation for the small apparent loss In stand between the two pre-harvest dates. The loss from the pre-harvest to final stand figures can be accounted for principally in small beets which were left in the field or went through the piler screens, and in beets of marketable size which were carelessly covered up or otherwise left in the field, or which fell from the trucks in delivery, as previously mentioned in connection with losses in yield. This difference in stand will probably hold fairly consistent from year to year.

The analysis of variance was made for all factory districts for both dates of sampling with the mean squares for yield being given in table 2. Analysis of variance was not made on percentage of sugar, since it is known that this character is considerably less variable than yield.

Table 2.—Mean squares for yield, all factory districts.

| | Mean Squares | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Classes | | | Between fields | | | Within fields | | |
| Factory district | df | 1st date | 2nd date | df | 1st date | 2nd date | df | 1st date | 2nd date |
| Bayard, Nebr. | 4 | 90.30 | 89.96 | 80 | 23.43 | 32.96 | 85 | 2.08 | 4.42 |
| Lyman, Nebr. | 4 | 178.44 | 145.26 | 45 | 30.18 | 39.90 | 50 | 1.23 | 0.68 |
| Mitchell, Nebr. | 4 | 45.05 | 68.38 | 40 | 29.76 | 37.34 | 45 | 6.88 | 12.39 |
| Brighton, Colo. | 4 | 50.96 | 53.88 | 56 | 28.56 | 25.50 | 60 | 0.57 | 6.21 |
| Brush, Colo. | 4 | 10.11 | 18.13 | 57 | 27.66 | 22.47 | 62 | 4.54 | 12.02 |
| Eaton, Colo. | 4 | 197.77 | 156.39 | 115 | 23.72 | 27.01 | 120 | 3.25 | 3.15 |
| Ft. Collins, Colo. | 4 | 50.02 | 58.87 | 85 | 15.74 | 18.52 | 90 | 2.17 | 1.51 |
| Ft. Lupton, Colo. | 4 | 144.79 | 173.64 | 60 | 31.57 | 32.62 | 65 | 2.71 | 2.09 |
| Ft. Morgan, Colo. | 4 | 110.94 | 179.43 | 94 | 24.52 | 32.50 | 99 | 3.60 | 1.29 |
| Greeley, Colo. | 4 | 81.52 | 54.05 | 95 | 25.56 | 20.92 | 100 | 4.66 | 4.15 |
| Longmont, Colo. | 4 | 108.61 | 60.60 | 85 | 23.23 | 18.88 | 90 | 5.08 | 4.90 |
| Loveland, Colo. | 4 | 75.26 | 130.82 | 70 | 24.42 | 31.30 | 75 | 3.64 | 1.04 |
| Ovid, Colo. | 4 | 48.65 | 90.83 | 110 | 22.07 | 27.89 | 115 | 4.48 | 3.26 |
| Sterling, Colo. | 4 | 103.41 | 85.54 | 95 | 27.63 | 34.50 | 100 | 3.61 | 4.20 |
| Windsor, Colo. | 4 | 61.96 | 143.18 | 75 | 18.56 | 31.33 | 80 | 2.75 | 2.45 |
| Lovell, Wyo. | 4 | 147.56 | 154.49 | 75 | 31.03 | 32.70 | 80 | 0.51 | 0.50 |
| General mean | — | 94.64 | 102.94 | — | 25.50 | 29.59 | — | 3.24 | 4.05 |

These results show a considerable difference in the effectiveness of the stratification or division into classes, with one exception, however, the mean squares for yield are large enough to justify the stratification. This fact is borne out by a general mean for "classes" for all districts compared with the mean for "between fields." In the case of Brush, Colorado, there was an actual loss in precision due to stratification, if we would arbitrarily assume that all samples had been randomly chosen without stratification, in which case the mean squares for "between fields" would have been 26.51 and 20.82 instead of 27.66 and 22.47 for the two respective dates.

A numerical estimate of the efficiency of stratification may be accomplished in the following manner, using Bayard, Nebraska, for example:

$$\frac{4\ (89.96)\ -\ 80\ (32.96)}{84} = 35.67.$$

The relative efficiency of the stratified to an unstratified sample is

$$\frac{35.67}{32.96} = 108.16 \text{ percent.}$$

If stratification had not been employed it would have been necessary to sample 8 percent more farms to maintain the same degree of accuracy as was obtained through stratification. The estimate, 35.67, is positively biased, but the bias is so small that for practical purposes it can be neglected, except in small samples.

Using this method of estimation, the relative efficiency of stratification for the various factor districts is as follows:

| | |
|---|---|
| Bayard, Nebraska | 108.16 percent |
| Lyman, Nebraska | 322.18 percent |
| Mitchell, Nebraska | 108.07 percent |
| Brighton, Colorado | 107.50 percent |
| Brush, Colorado | 98.75 percent |
| Eaton, Colorado | 115.08 percent |
| Fort Collins. Colorado | 109.09 percent |
| Fort Lupton, Colorado | 127.01 percent |
| Fort Morgan, Colorado | 118.46 percent |
| Greeley, Colorado | 104.09 percent |
| Lougmont, Colorado | 112.08 percent |
| JLoveland, Colorado | 118.18 percent |
| Ovid. Colorado | 104.09 percent |
| Sterling[1], Colorado | 105.97 percent |
| Windsor, Colorado | 118.07 percent |
| Lovell, Wyoming | 11S.50 percent |
| Mean | 112.25 percent |

As a mean of all factory districts, stratification resulted in an increase of 12.25 percent in relative efficiency.

A comparison of the variation "between fields" and "within fields' shows that very little was gained by taking 2 samples per field since the mean squares for "within fields" were generally much smaller than for "between fields." If we consider the two variances for the general mean of the later sampling date, making the assumptions that the same variation exists between other fields as between those sampled, and that the sampling variance would be doubled with 1 sample per field instead of 2, would have (29.59—4.05) + (2) (4.05) or 33.64 for the variance "between fields." This is an increase of 14 percent, which means that the number of fields would *need* to be increased only 14 percent to secure the same precision with 1 sample per field as was obtained in this study using 2 samples per field. Or, stating it another way, 1 sample from each of 114 fields will result in the same degree of accuracy as 2 samples from each of 100 fields. At the same time the number of samples would be reduced from 200 to 114, or 86 samples. In other words, we would need to sample, on the basis of 2 samples per field 75.4 percent as many fields as with 1 sample per field but with 1 sample per field only 57.0 percent as many total samples as with 2 samples per field, maintaining the same level of precision.

Another striking fact apparent from table 2 is the generally larger variance "between fields" for the second date of sampling. Three factories, Brighton, Brush, and Longmont, Colorado, are the only ones showing a smaller variance for the later date. This larger variance is to be expected, especially in a year characterized by soil moisture and temperatures favorable to late-fall development, such able .1.—Number of fields sampled with calculated precision, and number uf fields necessary for arbitrarily chosen fiducial limits for each factory district.

| Factory district | No. of fields sampled | Sampling precision ± tons per acre | Number of fields necessary with 1 or 2 samples each for fiducial limits (a of: X̄ ± 0.5 T, 1 | X̄ ± 0.5 T, 2 | X̄ ± 1.0 T, 1 | X̄ ± 1.0 T, 2 | X̄ ± 1.5 T, 1 | X̄ ± 1.5 T, 2 | X̄ ± 2.0 T, 1 | X̄ ± 2.0 T, 2 | X̄ ± 2.5 T, 1 | X̄ ± 2.5 T, 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | September 22 to 29 data | | | | | | | | | |
| Bayard, Nebraska | 85 | 1.62 | 1010 | 801 | 258 | 222 | 112 | 99 | 63 | 56 | 40 | 36 |
| Lyman, Nebraska | 50 | 2.35 | 1116 | 1110 | 279 | 275 | 124 | 122 | 70 | 69 | 45 | 44 |
| Mitchell, Nebraska | 45 | 2.40 | 1378 | 1035 | 345 | 259 | 153 | 115 | 86 | 65 | 52 | 39 |
| Brighton, Colorado | 61 | 1.71 | 869 | 699 | 217 | 175 | 97 | 78 | 54 | 44 | 35 | 28 |
| Brush, Colorado | 62 | 1.57 | 942 | 614 | 236 | 154 | 105 | 68 | 59 | 38 | 38 | 25 |
| Eaton, Colorado | 120 | 1.24 | 824 | 739 | 206 | 185 | 92 | 82 | 51 | 46 | 33 | 30 |
| Ft. Collins, Colo. | 00 | 1.19 | 548 | 508 | 137 | 127 | 61 | 56 | 34 | 32 | 22 | 20 |
| Ft. Lupton, Colo. | 65 | 1.85 | 062 | 889 | 241 | 222 | 107 | 99 | 60 | 56 | 38 | 36 |
| Ft. Morgan, Colo. | 99 | 1.49 | 009 | 874 | 227 | 218 | 101 | 97 | 57 | 55 | 36 | 35 |
| Greeley, Colorado | 100 | 1.35 | 835 | 724 | 209 | 181 | 93 | 80 | 52 | 45 | 33 | 29 |
| Longmont, Colorado | 90 | 1.22 | 641 | 509 | 160 | 127 | 71 | 57 | 40 | 32 | 26 | 20 |
| Loveland, Colorado | 75 | 1.67 | 878 | 850 | 230 | 212 | 98 | 94 | 55 | 53 | 35 | 34 |
| Ovid, Colorado | 115 | 1.28 | 835 | 748 | 209 | 187 | 93 | 83 | 52 | 47 | 33 | 30 |
| Sterling, Colorado | 100 | 1.52 | 1040 | 928 | 260 | 232 | 116 | 103 | 65 | 58 | 42 | 37 |
| Windsor, Colorado | 80 | 1.63 | 914 | 848 | 228 | 212 | 102 | 94 | 57 | 53 | 37 | 34 |
| Lovell, Wyoming | 80 | 1.66 | 898 | 885 | 225 | 221 | 100 | 98 | 53 | 58 | 36 | 35 |
| Mean number of fields | 82 | — | 912 | 803 | 228 | 201 | 102 | 89 | 57 | 50 | 36 | 32 |
| Mean number of acres (b for each sample to be taken | 100 | — | 9 | 10 | 36 | 41 | 90 | 92 | 144 | 164 | 228 | 256 |
| | | | September 2 to 8 data | | | | | | | | | |
| Mean number of fileds | 82 | — | 779 | 601 | 195 | 173 | 87 | 77 | 49 | 43 | 31 | 27 |
| Mean number of acres (b for each sample to be taken | | | | | | | | | | | 205 | 304 |

f t

Calculated from the formula N = V ( (M — X) ) N = No. of fields, V = Variance,
t for the 5% point, with (M — X̄) being the fiducial precision or the difference between the population mean and the sample mean.
Using 1 sample per 100 acres as a basis.

as obtained in these areas in the fall of 1941. Under these conditions some fields which have been largely depleted of available soil fertility will show early signs of so-called maturity, while others better sup plied with fertility will continue to show vigorous later growth, thus increasing the spread between the low and high-yielding fields.

The precision of the results obtained, together with the number of fields necessary to attain fiducial limits of from X=0.5 ton per acre to X=.2.5 tons per acre, are presented in table 3.

This table is presented for the later sampling, since it is evident from the mean squares in table 2 that more variation generally exists at this time than for the earlier sampling. Future consideration of rates of sampling should be based on the data for the later sampling. The means for the first sampling date are presented merely for purposes of comparison. Considering all factory districts, the average sampling precision of these tests was slightly more than X=1.5 tons per acre.

From the experience gained in conducting this test there are certain changes which might be made advantageously.

1. A geographic stratification by fieldman territories might be more practical than the one used which was based on 3-year previous yield records.

2. If contracts smaller than 5 acres in size are included, it is recommended that a sub-stratification based on contract acreage be made within geographic classes. This will have the effect of weighting the results on the basis of acreage.

3. Take but one sample per field, and determine the number of fields to sample based on the precision level desired.

## Summary and Conclusions

This study was conducted using random-sampling technique, with 2 arbitrarily chosen samples per farm and with the farms chosen within strata based on 3-year previous yield records. One farm per each 100 acres in commercial beets was the basis used for sampling, the study being conducted for all 12 factory districts in Colorado, 3 of the Nebraska factory districts, and the Lovell, Wyoming, district. The most significant findings and conclusions are as follows:

1. The farms chosen proved to be almost perfect samples for the Colorado and Nebraska districts and also for the Lovell factory district as was indicated by the fact that the average harvested stands, yields, and percentage of sugar for the farms sampled almost duplicated the corresponding final factory averages.

2. The results indicate for this year a large increase in percentage of sugar and striking losses in final delivered-per-acre yields from the percentage of sugar and total tonnage indicated by the pre-harvest samples. Sources of probable loss are explained both for tonnage and for stand.

3. The stratification, which was based on previous 3-year yield records, proved effective for all but one factory district.

4. The variance between the 2 samples per field was small, and it is shown that by increasing the number of fields 14 percent, the total number of samples could be reduced to 75.4 percent of the number taken this year with the same level of precision resulting. On the basis of 1 sample per field (or farm), 1 field per 102 acres would result in a sampling precision of =1.5 tons per acre, this being based on the 5-percent point level.

5. The necessary number of fields with 1 or 2 samples per field was calculated for each factory district for sampling precision levels of rtO.5 to ±2.5 tons per acre.

6. Suggestions are made for changes in future studies to include: (a) Geographic stratification, (b) using a sub-stratification based on contract acreage if contracts or farms of less than 5 acres are to be sampled, and (c) taking only 1 sample per field.

7. This method of sampling appears to give an accurate estimate of the condition of the crop at the time of sampling. The prediction from the last sampling date to final harvest figures should become increasingly precise, as accurate pre-harvest estimates are accumulated over a period of years, and the source of errors in this estimate become more fully recognized.

## Literature Cited

1. King, Arnold J., and Jebe, Emil H.
1940. An experiment in pre-harvest sampling of wheat fields. Iowa Agri. Exp. Sta. Res. Bui. 273.

2. King, Arnold J., and McCarty, Dale E.
1941. Application of sampling to agricultural statistics with emphasis on stratified samples. The Journal of Marketing. April, pp. 462-474.