An Experiment in Pre-Harvest Sampling of Wheat Fields

By Arnold J. King and Emil H. Jebe

AGRICULTURAL EXPERIMENT STATION
IOWA STATE COLLEGE OF AGRICULTURE
AND MECHANIC ARTS

STATISTICAL SECTION

AGRICULTURAL MARKETING SERVICE
UNITED STATES DEPARTMENT OF AGRICULTURE
WORK PROJECTS ADMINISTRATION
NEW YORK CITY
Cooperating

RESEARCH BULLETIN 273
AUGUST, 1940
AMES, IOWA

August, 1940

Research Bulletin 273

An Experiment in Pre-Harvest Sampling of Wheat Fields

By Arnold J. King and Emil H. Jebe

AGRICULTURAL EXPERIMENT STATION
IOWA STATE COLLEGE OF AGRICULTURE
AND MECHANIC ARTS

STATISTICAL SECTION

AGRICULTURAL MARKETING SERVICE
UNITED STATES DEPARTMENT OF AGRICULTURE
WORK PROJECTS ADMINISTRATION
NEW YORK CITY
Cooperating

AMES, IOWA

CONTENTS

SUMMARY

This is the report of a preliminary investigation concerning the possibilities of estimating and forecasting wheat yields by an objective method of sampling the wheat crop as produced under farm conditions. The report gives a description of 1 year's work, using methods that are still in the experimental stage.

These investigations are based upon a sample taken in the eastern half of North Dakota just prior to the 1938 harvest.

The objectives of the sampling were to investigate:

1. The practicability of a route method of sampling the wheat crop to estimate and forecast yields per acre.
2. The amount of information that might be gained by the use of stratification in the sampling—by geographical division of the area sampled and the identification of the varieties in the samples taken.
3. The nature of the variation of yield among fields and the variation within fields and their relative magnitudes.
4. The kind of crop counts and measurements that may give the best basis for estimating and forecasting wheat yields, estimation being defined as the determination of the yield just prior to, or at, harvest time, while forecasting is considered as predicting the yield at some time previous to harvest.
5. The extent of the bias in the sampling.

An Experiment in Pre-Harvest Sampling of Wheat Fields[1]

By Arnold J. King[2] and Emil H. Jebe[3]

There has developed a demand on the part of millers, farmers, experiment stations and several government agencies for more accurate forecasts of yield per acre during the growing season, and especially for more accurate, timely estimates of acreage, yield and quality of the several varieties of wheat grown in the more important wheat-producing states. These estimates are needed at harvest and prior to marketing. The farmers need such information in planning their sales of the current crop, as well as in planning their future farm operations. The millers need to locate the kind of wheat best fitted for their needs. Those interested in improving the quality of the wheat crop are desirous of having such estimates made because of the incentive furnished for the growing of the varieties having the better milling and baking qualities. One of the first needs of experiment stations interested in developing methods for disposing of surplus wheat is accurate information on the quality and production of wheat by varieties for areas within the important wheat-producing states.

---

[1] The sampling was conducted by the Agricultural Marketing Service, U.S.D.A., the field work being done by A. J. King and J. E. Pallesen. The samples were forwarded to New York City, where threshing, counts and measurements were made with the aid of personnel provided by the Work Projects Administration. The authors directed the statistical analysis in collaboration with the staff of the Statistical Laboratory, Iowa State College. The authors are especially indebted to G. W. Snedecor, W. G. Cochran and Gertrude M. Cox for their advice, suggestions and criticism; also to members of the Agricultural Marketing Service for suggestions regarding the presentations of the material and especially to C. F. Sarle and W. F. Callander of the latter agency, under whose supervision the study was made possible.

This study was conducted by the Bureau of Agricultural Economics in 1938-39. With the establishment of the Agricultural Marketing Service on July 1, 1939, the work (together with the personnel named) was included among the functions transferred to the latter agency of the Department of Agriculture.

Project 611 of the Iowa Agricultural Experiment Station.

[2] Agricultural Statistician in the Agricultural Marketing Service, U.S.D.A., assigned to the Bankhead-Jones research project, Research in the Statistics of Agriculture and Associated Statistical Theory, conducted in cooperation with Iowa State College.

[3] Research Graduate Assistant, Statistical Section, Iowa Agricultural Experiment Station.

The chief source of information about the wheat crop in the United States is the reports issued by the Crop Reporting Board of the United States Department of Agriculture. These reports, issued throughout the growing season near the first of each month, are confined in scope to forecasts and estimates of production by states. These estimates are based upon questionnaires distributed by mail, and although they are of material help to the wheat industry and the agricultural programs, they are inadequate for present-day demands. The reports lack in timeliness, because they are issued only once a month and cover states as a whole. Since the date of harvest is seldom uniform over a state, the reports do not closely follow the harvest. The estimates are not made by varieties, and they do not give any information about the quality of the crop. The method used for collecting the data lacks in objectivity, and this, together with the fact that the number of reports is small, does not, in many cases, permit the making of estimates and forecasts for areas smaller than the state.

For several years it has been thought that at a small cost at harvest it would be possible to obtain a sample of heads in an objective manner of sufficient number to estimate the yield per acre and quality of the wheat crop by varieties in groups of counties in the more important wheat-producing states of Texas, Oklahoma, Kansas, Colorado, Nebraska, South Dakota, Minnesota, North Dakota and Montana. It was also thought that such a sample would give a basis for determining to what extent structural counts of the attributes of yield may be useful in the development of objective methods of forecasting yield per acre in advance of harvest. The desire to test out these assumptions led to this investigation, employing an objective sampling process applied to the 1938 wheat crop in the eastern half of North Dakota just prior to harvest.

The most important previous work in this field has been carried on by English statisticians. One American investigation, however, is worthy of note. Quisenberry (1) (1926) reports a sampling project undertaken in six counties distributed among four states. Two hundred seventy-three fields were sampled in this study. Due to uncontrolled factors the sample yields and the yields reported by the farmers were not very highly correlated. The English work originated at Rothamsted has been frequently discussed by Yates and Cochran (2) (3) (4) (5) (6). Bias and efficiency of the sampling unit have been treated by Yates and Zacopanay (3). The problem of determining the mean yield of a geographic area, such as a crop-reporting district or state, has received the attention of the English statisticians.

The progress of the English workers and their understanding of the problem, the estimation of yield from weighed samples of grain, may be indicated by quotations from their writings. Yates (4): "In view of the large variability in the yields of fields in a given district, it is not to be expected that the yields of the experimental crops should bear any close relation to the mean yield of a district.... The role of the experimental plots is to indicate the most useful observations." Cochran (6): ". . . the estimation of the mean yield of a country by weighings at harvest is mainly a question of organization. Here the use of sampling promises to be helpful, indeed essential, both for determining the yield of a field and of a district. Preliminary research would be needed to develop a good sampling technique, to train observers in its use and to assess the amount of time and labor required to estimate the mean yield of a country with a given degree of accuracy.

"The provision of unbiased estimates of the yield of a crop before harvest and at harvest does not settle the whole problem of crop estimation. Losses occur in the cartage and storage of crops, and in some cases these are quite large, so that the yield at harvest time is an over-estimate of the total crop which ultimately reaches the market. Allowance for losses after harvest is a separate problem, but it should not prove insuperable."

SAMPLING PROCEDURE

In the planning of the survey, several practical questions presented themselves. Would it be possible to secure mature samples uniformly over the area to be sampled? Uneven ripening of the grain would make uniform sampling very difficult. How should the fields be selected for sampling? An unbiased sample would be obtained from fields selected at random, but the impracticability of such selection made route sampling necessary with the retention of as large an element of randomness in the sampling as possible. The extent of the bias in route sampling is unknown. Previous work by the Agricultural Marketing Service in using routes for crop metering acreages indicated that routes could be used satisfactorily. Routes should be laid out to follow any gradient, such as a yield gradient, which may be present in the area being sampled. Thus, the full extent of the variation in the area sampled would be included, and the best estimate of yield would be obtained. The size of the sampling unit and the number of units to be taken per field presented a further problem. Various investigators (3) (7) had used, with considerable success, 1/10,000 of an acre as a unit in sampling experimental plots.

Accordingly, this unit was selected, two of them to be taken from each field sampled.

The measurements and counts in the samples decided upon were:

1. Average height of plants
2. Number of heads
3. Average length of heads
4. Number of sterile spikelets
5. Number of fertile spikelets
6. Average number of kernels per spikelet
7. Average number of kernels per head
8. Number of kernels in sample
9. Weight of 200 kernels in sample
10. Weight of the sample
11. Variety identification

The field party for the sampling consisted of two men traveling in one car. The car was equipped with a crop meter to measure the frontage of wheat, and the odometer was used to determine the geographic miles driven within a county or district.

Three evenly-spaced routes running north and south and three running east and west were laid out in each of the five crop-reporting districts which comprise most of the eastern half of North Dakota. Thus, the routes formed a grid-like pattern over each district. The county 10-year average yields in North Dakota indicated a yield gradient running east and west. Greater efficiency would probably be obtained in most years if only east and west routes were used. On the other hand, it is possible that in some years the gradient could run north and south. In that event, east and west routes would be very inefficient. Since the variability gradient was not known definitely, the grid pattern was chosen for the routes in the first survey. The routes were made as straight as possible, and thus, the geographic miles driven by the car in each district were approximately in the same proportion to the geographic area of the district for all districts. In this way, each district tended to receive its proper weight in the sampling.

Selection of the fields to be sampled was determined by the crop meter. After each 2 miles of wheat frontage registered by the crop meter on one side of the road along the route, the next field was sampled. The crop meter provided an additional control on the distribution of the sampling in that not only land area, but area in wheat, controlled the sampling.

Upon arriving at a field so selected, the samplers made an estimate of its frontage. Then two points at which the field was to be entered were selected at random (by the use of a set of random numbers and some unit of frontage measure such as fence posts or telephone poles). In order to eliminate the "border effect" and at the same time to reduce the amount of walking, the area sampled was restricted to a strip lying

between 20 and 120 paces from the edge of the field. To determine the points within the field at which the samples were to be taken, two more random numbers between one and one hundred were selected. These numbers determined the number of paces (beyond the 20-pace border) which the sampler was to walk into the field from the point of entry. This method of selecting the samples made the sampling as nearly random as the definition of the population (a specific strip within a field lying along the route sampled) would permit.

Each sample consisted of an area 24" by 26.14", which is approximately 1/10,000 of an acre. Each sampler used a U-shaped bar or hoop to measure the sampling unit. The bar was placed perpendicular to the drill rows and, since most of the wheat was drilled in 6-inch rows, 4 contiguous drill rows were included in each sample.

Following the selection of the sampling unit and the placing of the hoop, measurement was made of the average height of the plants in the unit. An estimate of the extent of grasshopper damage was recorded. Variety identification was made in the field whenever possible. The grain within the hoop was then clipped and placed in a labeled envelope. The wheat mileage from the crop meter and the geographic mileage from the odometer were recorded for each sample. With the routes marked on a highway map it was then possible to determine the approximate geographic location of each sample.

The samples were packed and shipped to New York City for the laboratory work. The laboratory work was done by W.P.A. personnel under the direction of an agronomist. The agronomist also checked the variety identification against prepared samples of the varieties grown in North Dakota which were obtained from the state experiment station. Then the threshing and counting were done to obtain the desired measurements of the samples.

SUMMARY OF THE DATA

The data secured and tabulated include the following parts.

Samples were taken in five crop-reporting districts consisting of 23 counties in eastern North Dakota. Two hundred thirty-two fields were sampled, two samples being taken from each field in the manner described above. Identification of the samples divided them into six varieties and one species, Durum. Samples of three varieties,[4] Ceres, Durum and That-

[4] Hereafter in this report reference is made to Durum as a variety for ease of presentation. Durum is a species of wheat, several varieties of which are grown in the area sampled. These varieties were not distinguished by the samplers.

cher, were obtained in sufficient number for statistical analysis. In addition to the 10 measurements and variety identification listed on page 628, the geographical location of the sample by county and district was determined. Particular field conditions noted by the sampler were also recorded.

VARIETY AND DISTRICT YIELDS

TABLE 1. NUMBER AND MEAN YIELD (BUSHELS PER ACRE*) OF WHEAT SAMPLES FROM SAMPLING SURVEY IN NORTH DAKOTA—1938.

Crop-reporting district	Ceres		Durum		Thatcher		Total	Mean
	Number samples	Yield	Number samples	Yield	Number samples	Yield		
3	36	10.3	134	15.8	46	18.2	216	15.4
5	12	7.8	12	11.8	14	14.2	38	11.4
6	36	10.2	8	14.9	92	19.8	136	17.0
8	4	3.4		‡	2	19.8	6	8.9
9	20	6.8	16	14.8	12	21.4	48	13.1
Total	108		170		166		444	
Mean		9.0		15.4		19.0		15.2

* The yield is based on the conversion of the number of grams of wheat in the 1/10,000-acre sample to bushels per acre.
‡ No Durum samples in district 8.

The distribution of the number of samples by district and variety shown in table 1 was determined by the particular sampling procedure adopted. This was explained in the section on sampling procedure. If the choice of the routes which were followed in the sampling gave a representative sample of the area, then the number of samples is approximately proportional to the acreage of wheat in the district. The number of samples taken was determined by the use of the crop meter. As indicated previously (p. 627), the experience of the Agricultural Marketing Service in using routes to meter acreages has given satisfactory results. Similarly, the number of samples of a variety tends to show the proportion of the wheat planted to that variety.

An examination of table 1 shows that Ceres is the lowest and Thatcher the highest yielder among the three most widely distributed varieties. Furthermore, this rank is uniform through all the districts. This may be a seasonal characteristic. Rust was exceptionally bad in parts of the area sampled in 1938, and there was considerable grasshopper damage. Among the varieties, Ceres was most affected. Thatcher, on the other hand, a new rust-resistant variety recently introduced in this area, suffered least from both rust and grasshopper damage. It is recognized that in another season the yield ranks of the varieties may change.

The district yields are shown in the right hand column of table 1. The small number of samples in district 8 makes the mean unreliable. The reasons for taking such a small number of samples in this district will be discussed later. (See pages 636 and 644.)

ANALYSIS OF VARIANCE OF YIELD

TABLE 2. ANALYSIS OF VARIANCE OF YIELD* WHEAT SAMPLING SURVEY—NORTH DAKOTA—1938.

Source of variation	Degrees of freedom	Sum of squares	Mean square
Samples from the same field	222	4220	19
Fields of the same variety in a district	208	13512	65
Varieties in the same district	9	6079	675
Districts	4	1430	358
Total	443	25241	

* Unit is bushels per acre.

SAMPLING VARIATION WITHIN FIELDS

A comparison of the variation in the same field with the variation among fields (mean square of 19 and 65, respectively) shows that sampling per field was adequate under 1938 conditions in the area sampled. This can be quite easily illustrated. Doubling the amount of sampling per field would reduce the sampling variance (or mean square, 19) by one-half. The total variance of fields would then be $(65 - 19) + 19/2 = 56$. This would reduce the number of fields which would have to be sampled in order to get the same accuracy by 14 percent $[100(1 - 56/65) = 14 \text{ percent}]$. That is, only 86 percent, or 190 fields, would be required if there were four samples per field. On the other hand, these four samples per field would make the total number of samples 760 instead of 444. That is, the slight reduction in number of fields to be sampled would be outweighed by the increase in labor resulting from handling almost twice as many samples.

It may be interesting to consider the opposite case—reducing the number of samples taken per field. Here it may be assumed that the sampling variance would be doubled. The variance of fields would then be $(65 - 19) + (2)(19) = 84$. The number of fields to be sampled would be increased by 29 percent, 287 fields being required to give the same accuracy as that attained in the actual sampling. However, though this change in sampling would reduce the number of samples to be handled from 444 to 287, it would require driving more miles or making more stops in the same distance. On the other hand, in taking one sample per field, one man per car might be allotted the same amount of territory. This would be a decided advantage in reducing the cost of the sampling through reduction of the field personnel required. However, the taking of one sample per field would result in the loss of information about the variation within the fields. This information is important for the statistical examination of the results. Furthermore, the proper distribution of the resources in sampling depends on the relative magnitudes of the variance among fields and within fields. Since this relation between the two variances may not be the same in other seasons or in other areas, it would seem advisable to continue taking two samples per field until more information is acquired about the variation among fields as compared with the variation within fields.

SAMPLING VARIATION AMONG FIELDS

Consideration of the variation among fields from another point of view makes it possible to estimate the number of fields which must be sampled to obtain a desired degree of accuracy. If the sample mean is to lie within a specified range of the true mean in 19 out of 20 cases, the number of fields which must be sampled for this degree of precision may be estimated from the formula, $M = \bar{x} \pm \frac{st}{\sqrt{n}}$, which gives the fiducial limits of the mean. Solving this formula for n gives $n = V\,[t/(M - \bar{x})]^2$, replacing s^2 by V, the variance. As an example, it may be desired to place the fiducial limits of the mean at $\bar{x} \pm 1$ bushel and to determine n for the probability .95. Since our estimate of the variance is based on 222 fields, we may use $t = 1.971$, and with $M - \bar{x} = \pm 1$, the formula becomes $n = 3.88V$.

However, before proceeding directly to the application of this formula to our data, it seems proper to indicate the alternatives for its use which depend on the data available. Further, the assumptions which are made in using the formula must be kept in mind. If the proportions of the varieties grown in each district are known, the appropriate sampling error mean square for use in the formula is 65, the variance (table 2) among fields of the same variety within a district. With $V = 65$, the formula gives $n = 253$, which is only about 14 percent more fields than the rate of sampling in the 1938 project. In this calculation it is assumed that increased sampling

will not change the variance among fields.

On the other hand, if the proportions of the varieties grown in the districts are not known, the mean square for computing n to be used in the formula is the variance among fields within districts which is $(13512 + 6079)/217 = 90$. Now, with $V = 90$, computing n gives 350 fields. This result is about 40 percent greater than that obtained above with the assumption of the proportions for the varieties being known.

In these calculations, as in any application of fiducial limits, it is implied that the estimation of the mean is without bias. As will be shown later (p. 643), there seems to be evidence of bias in some of the sampling. If bias can be removed, then, as the knowledge of field to field variation is increased in future sampling work, such calculations as the above can be used with increased confidence.

A further assumption which must be made is that route sampling gives a proper estimate of the sampling error. Route sampling does not give a random selection of fields, but practical difficulties make random choice impossible. A poor choice of routes may under or overestimate the sampling error, yet the desirable features of route sampling which will be pointed out later (p. 646) and the difficulties of random choice, leave route samplings as the preferable method. Selection of the routes must be made to avoid any known bias.

VARIETY AND DISTRICT VARIATION AND STRATIFICATION

The differences among the variety means in the districts as shown in table 1 are highly significant. This is evident from the analysis of variance shown in table 2 (mean squares of 675 and 65, respectively, for varieties in the same district and fields of the same variety in a district). As indicated previously, this may be a peculiarity of the season. However, it is quite possible that the characteristics of these varieties, or other varieties which may be introduced, are such that in no season will the varieties tend to yield the same within the same district or over all districts. An estimate of the gain due to variety stratification may be obtained from the previous paragraph (p. 632) on fiducial limits. Without variety identification in order to determine the proportions grown of the varieties, about 40 percent more fields would need to be sampled in order to secure the same accuracy in estimation. This statement is based on the 1938 results in North Dakota. Should this result be fairly consistent from year to year, variety identification becomes a requirement in the sampling. Certainly, until our knowledge is greatly expanded, this identification must be continued.

When the problem of estimating total production is attacked, increased sampling for variety identification alone may be found to add considerably to the sampling information. Extra stops can be made easily along the route in order to collect the small amount of wheat necessary for identification in the laboratory. The increased sampling would allow the estimation of the proportions of the varieties more accurately. This information would probably be a worthwhile addition to the knowledge of the wheat as produced under farm conditions.

The variation among districts is not independently evaluated in table 2, where the apparent district variation is a composite of variety, field and the true district variation. It is of some interest to obtain an estimate of the district variation after allowing for the variety effect. This analysis also permits an estimate of the variety x district interaction. The appropriate analysis yielded the results shown in table 3.

TABLE 3. ANALYSIS OF VARIANCE OF YIELD
WHEAT SAMPLING SURVEY—NORTH DAKOTA—1938.

Source of variation	Degrees of freedom	Sum of squares	Mean square
Variety (without considering district)	2	6492	3246
District (after allowing for variety)	4	621	155
Variety x district interaction	7*	396	57
Fields of same variety in a district	208	13512	65

* One of the varieties is missing in district 8. Thus, there are only 7 degrees of freedom for the interaction.

The district mean square (155) in table 3 is somewhat higher than the mean square (65) between fields of the same variety in the same district shown in table 2, but not significantly so. This indicates that the real differences between districts were not sufficiently large and consistent to show up definitely in a sample of this size taken in the 1938 season in North Dakota. A fair proportion of the apparent differences among the districts, as shown by the district means in table 1, must be attributed to the varietal differences. By referring to table 1 it may be noted that the district with the highest yield, district 6, has a very large proportion (92 samples of a total of 136 in the district) of Thatcher, the highest yielding variety. The other districts with lower yields contain smaller proportions of Thatcher.

A further point may be noted here. A comparison of the variety x district interaction mean square (57) with the mean

square (65) for variation among fields of the same variety in a district shows no indication of a real interaction. That is, the varieties tended to perform the same relative to one another over all the districts sampled in North Dakota in 1938. This, of course, may be a characteristic of the 1938 season. Future sampling may show quite different effects.

Estimation of the reduction in sampling error due to stratification by districts is complicated.[*] The result depends on whether stratification by varieties also is being employed. If stratification is by districts, but not by varieties, the estimated sampling error is derived from the total mean square between fields within districts, which from table 2 is $(13512 + 6079)/217 \doteq 90$. The corresponding mean square without the use of district stratification is a weighted mean of this figure and 358, the mean square between districts in table 2. On the other hand, if stratification is both by districts and varieties, the sampling error mean square is 65, the variation between fields of the same variety in a district. The comparable sampling error for stratification by varieties, but not by districts, is a weighted mean of 65 and 155, the variation between districts after allowing for varietal effects.

In both cases it appears on further investigation that the gain due to stratification by districts was small. It is not deemed wise, however, at the present writing, to abandon this stratification. Several reasons may be advanced for retaining this feature — geographical stratification — in the sampling. Present information available is based on only one season's results. Such stratification is a matter of convenience, the crop-reporting district being the present geographical unit upon which the Agricultural Marketing Service bases its estimates of yield and production. The use of a larger unit than the crop-reporting district is hardly feasible. Recording the geographical location of the field from which the sample is taken for any stratum which may be chosen, such as county or district, is easily done.

Furthermore, the variation among districts was probably underestimated in North Dakota in 1938. The unreliability of the sample mean for district 8 was indicated above. (See page 631.) In fact, for the reasons outlined below, this mean (8.9) is probably too high. The samples were to be taken from "harvested acres" as defined by the Crop Reporting Board in its estimates. Thus, the sampling results would then be comparable to the regular estimates. Difficulties which arose in

* The authors are indebted to W. G. Cochran for this discussion on the reduction of the sampling error by stratification.

the definition of this term, "harvested acres," will be discussed later. (See page 644.) Crop conditions were very adverse in district 8 in 1938. Rust and drouth together lowered the yield or destroyed the crop almost entirely. Many low-yielding fields were not sampled, since it was assumed that they would not be harvested. Consequently, since most of the low yields were found in district 8, the variation among districts was greater than the sampling results indicate. Stratification by crop-reporting districts in the 1938 sampling has furnished information on the proportions of the varieties grown and where these proportions are grown. The performance of the varieties in one season in several districts can be compared. With improvement in the sampling procedure to include stratification by counties, that is, keeping the length of the route in each county proportional to the area of the county and the proportion the same for all counties, a further worthwhile gain in information may be obtained.

ESTIMATING AND FORECASTING BASED ON SAMPLE COUNTS AND MEASUREMENTS

A fourth objective of this project was to investigate the possibilities: 1. Of estimating the yield per acre of wheat just prior to harvest time on the basis of plant characteristics ascertainable in the field, and 2. of forecasting the yield earlier in the season from plant characteristics that can be measured some time before harvest. These two problems are of practical importance. If certain readily observable plant characteristics are highly correlated with yield per acre, the crop estimator may ultimately be able to make observations and estimate the yield directly from the knowledge of these characteristics. The earlier these observations can be made in the wheat fields, the greater will be their value for predicting the crop. The development of such a method from an objective standpoint may lead to great improvement in the present methods of forecasting and estimation. Perhaps much of the laboratory work of this preliminary study might be unnecessary.

During the past two decades a number of statistical investigations relating various factors to the yield of wheat have been made. These have been undertaken by plant breeders and agronomists interested in developing new and improved varieties and increasing the yield of wheat. Numerous characters have been correlated with yield by these workers in search of leads which might aid their research.

Sprague (10) found a significant relation between yield and average number of spikes per unit area: Hayes, Aamodt and

Stevenson (11) in their correlation studies report date of heading, height of plants and plumpness of grain as important factors in relation to yielding ability of spring wheat. Bridgford and Hayes (12) (13) in their investigations also showed date of heading and height to be positively correlated with yield. Immer and Ausemus (14) found plumpness of grain to be closely associated with yielding ability. Laude (15) presented graphs covering a 6-year period showing the relation of number of heads per unit area, test weight of grain and kernel weight to yield of wheat. In Quisenberry's sampling study (1) multiple correlations of the sample yield with three characters, number of heads, weight of 1,000 kernels and number of kernels per head, were high.

More recently the English statisticians quoted previously (p. 626) have studied the problem as approached in this experiment. The English investigations have shown plant number and shoot height to be significantly associated with yield. Yates (4) writes: ". . . forecasting based only on the detailed study of a few experimental plots, though it may predict the yields of these plots with great exactitude, is not likely to be very successful in predicting the mean yield of a district. The role of the experimental plots is to indicate the most useful observations. The prediction of the average yield of a district only can be undertaken by taking measurements on commercial crops.

"It should also be emphasized that such measurements would have to be taken for several years before forecasting of any kind could be attempted, for it may well be that a forecasting formula that gives a good result for the experimental plots will require modification before it can be applied to commercial fields. To mention only one disturbing factor, differences in varieties will clearly introduce complications."

ESTIMATING YIELD FROM PLANT CHARACTERISTICS

In making this study it was assumed that each of the nine measurements of the samples would have some association with yield per acre of wheat. As a first step in testing this assumption, the relation of the plant characteristics to yield was examined by the method of multiple regression. Each variety was analyzed separately because of the large varietal differences shown in table 1. From each regression the variation among fields was removed by the methods of multiple covariance. The resulting regressions were based only on the relations among the measured factors existing within the same fiield from which each pair of samples was taken. The value of R^2 (square of the multiple correlation coefficient) was large for each variety (.71—Ceres, .90—Durum, .98—Thatcher) indicating that the yield of a sample is quite closely associated with the variables contained in the complex of nine. The value of R^2 for Ceres was lower than that for the other two varieties, which may be explained by the fact that this variety suffered most from rust and grasshopper damage in 1938. In Ceres, the weight of 200 kernels had very little relation to yield. Examination of the data revealed that this variable was practically constant for all samples of Ceres. Variation in yield in Ceres was almost entirely due to differences in the number of kernels per sample. Number of kernels per sample also contributed the most information for Durum and Thatcher, yet the weight of 200 kernels added considerably to the knowledge of yield.

Durum alone presented some peculiar relationships. The average number of kernels per spikelet showed a negative relation to yield while number of kernels per head showed a positive relation. The weight of 200 kernels contributed more information for Durum than for the other varieties. These facts bear out an observation of the samplers—well-filled heads of Durum with large kernels were the best yielders. Since the Durum wheats were not distinguished as to variety, it is not known if this was a characteristic of a particular Durum wheat. Varietal classification of Durum in future sampling may yield some information on these points.

In the complex of nine variables studied, the number of kernels per sample contributed the most information for all varieties. Regardless of this uniformly close relationship of number of kernels per sample to yield, this is a measurement that is not easily determined in the field. Of great importance to the crop estimator is the ease of making observations. The number of kernels in a sample can be determined only by actual harvesting and threshing. If the sample has to be threshed the grain can be weighed and the yield is then known without recourse to any regression.

Since number of kernels per sample is not a convenient variate for determining the yield, the other variables in the complex may be examined as possible sources of the same information in the absence of number of kernels per sample. Number of heads and length of heads will give some indication of the number of kernels. The only other variable in the group which can be measured easily is height of grain in the sample. Using these three variables and recomputing the regressions gave the results shown in table 4.

The number of heads per sample now contributes the most information in the absence of number of kernels per sample.

TABLE 4. STANDARD PARTIAL REGRESSION COEFFICIENTS ON YIELD FOR THREE VARIABLES BY VARIETY USING THE FIELD AS A UNIT FROM THE WHEAT SAMPLING SURVEY IN NORTH DAKOTA IN 1938.

Variable	Ceres	Durum	Thatcher
Number of heads	.61	.63	.90
Average height of grain in sample	.17	.23	.14
Average length of heads	.26	.27	.27
R^2	.64	.60	.88

Length of head is next in importance. Height contributes the least information. English investigators (4) have found height most highly correlated with yield. It may be that this is a characteristic of regions where rainfall is plentiful. However, in the Great Plains wheat belt where rainfall is often deficient, height may not prove to be so important as an indicator of yield. The value of R^2 is somewhat smaller than before for each variety regression. Durum, particularly, is not so well estimated as when nine variables were used. This may be explained by the fact which was observed that neither height nor number of heads per sample were very closely related to the yield of Durum. (See page 638.)

These results (table 4) give some hope that further study of the relations may be worthwhile. The smallest value of the square of the coefficient of multiple correlation is .60. It is realized that this study is based on samples collected in only one year, 1938. Analysis of data collected in another year may show different results. For these regressions to become useful they must be extended over a number of seasons. Furthermore, the samples for this study were taken in only a small part of the wheat belt, five crop-reporting districts in eastern North Dakota. Other areas may show quite different relations. In a year of severe crop damage, stem (black) rust for example, the yield per acre is reduced to almost zero. Under such conditions, plant characteristics would have little, if any, correlation with yield. In fact, only the weight of the grain in a sample can be depended upon under highly adverse conditions. However, it is hoped that the analysis of data from future sampling will furnish more exact information about criteria which are related to yield and the effect of season on these relationships.

FORECASTING YIELD FROM PLANT CHARACTERISTICS

The wheat survey of 1938 did not provide any data for direct use in investigating the possibilities of forecasting the yield of wheat. However, the preceding discussion indicates the nature of the problem. For purposes of forecasting, one must select plant characteristics which can be measured some time before harvest. In addition, as in estimation, it is advantageous to observe those attributes which may be measured easily in the field. The earlier in the season these measurements can be made, the greater is their utility. Stand is perhaps the most important objective measurement of value before the wheat is headed. After heading, the variates listed in table 4 can be measured. Although the relationship of these variables to yield at the time the grain heads may not be the same as at harvest time, yet since these measurements can be made several weeks in advance of harvest, there is a possibility of basing a prediction upon them which could be issued as a forecast of yield. A study of observations on wheat taken at various times before harvest thus presents an attractive field for exploration. If measurements of the variates listed in table 4, when taken earlier in the season, should give comparable results, then number of heads may be used as a predictor in forecasting yield.

ESTIMATION USING THE COUNTY AS A UNIT OF AREA

Thus far, an attempt has been made to discover the relations existing between each of several plant characteristics and the yield of wheat. These results might be applied to estimating the yield of wheat for the field from which the samples were obtained. An illustration of how these results, if they should prove consistent over time, may be applied by the Agricultural Marketing Service in its work would further indicate their utility. In view of the fact that much of the work of the Department of Agriculture is on the county basis, this unit of area was selected for the following analysis. Here, the sample yield per acre of a county is used instead of that of the individual field. This procedure seems useful in that the county is the administrative unit for which quotas and estimates are prepared by the Agricultural Adjustment Administration and the Soil Conservation Service.

The variables selected for this analysis are listed in table 5. The county sums for each variable in each variety were accumulated. These sums were then related to the county yield. The resulting regressions are those between counties. These include the variation among counties in contrast to the previous regressions which contained only the variation existing between the two samples in the field from which the samples were taken. The results of this method of computing the regressions are presented in table 5.

TABLE 5. MEANS AND MULTIPLE REGRESSION ANALYSIS OF THREE VARIETIES OF WHEAT WITH THE COUNTY AS A UNIT OF AREA FROM THE WHEAT SAMPLING SURVEY IN NORTH DAKOTA IN 1938.

	Ceres	Durum	Thatcher
Means			
Yield (bushels per acre)	9.05	15.39	19.00
Number of heads in sample	83.00	67.00	129.00
Average height of grain (inches)	27.50	32.10	29.50
Average length of heads (inches)	2.80	2.19	2.42
Standard partial regression coefficients of yield on:			
Number of heads in sample	.80	.41	.60
Average height of grain	.00	.71	.36
Average length of heads	.07	—.37	.22
R^2	.70	.74	.91
Standard error of estimates (bushels) of the mean yield per acre for a county	1.14	3.07	0.98

As might be expected, the shift in the unit of area in the analysis brings out relations among the variables quite different from those determined before. By comparison of table 5 with table 4 it is seen that R^2 has changed little. But several differences may be noted in the values of the standard partial regression coefficients. Height now contributes the most information for Durum, and more information than length of head for Thatcher. This indicates that for counties as a whole the counties with the taller wheat had the higher yields, while height did not have a high relation to yield within the same field from which the two samples were taken. Length of head is negative in its relationship to yield for Durum. This further substantiates the observation of the samplers that short, plump heads of Durum contain more wheat. As mentioned previously (p. 638) this may be a varietal characteristic of one of the Durums. Number of heads contributes almost all the information for Ceres. This again bears out previous observations concerning Ceres for the 1938 season in North Dakota. Consideration of the values of the standard errors of estimate indicates that the fiducial limits which may be placed on the estimated yields for the counties are rather wide. This might be expected from the smaller number of samples in some of the counties.

The results of this preliminary investigation of estimation can be considered only as indicating possible results which may be obtained from future sampling. Perhaps the accumulation of information over time will point out definitely the observations of the wheat plant which should be taken to estimate yield. Only after considerable information has been accumulated on the regression of these or other variates on yield, with consistent results over time, may the hypothesis be set up that a "true" regression exists. Then this regression may perhaps be used for making estimates and the placing of desired fiducial limits on the estimates.

SOME PROBLEMS CONNECTED WITH SAMPLING

The difficulties encountered in securing the wheat samples may be classified under two general heads. The first of these is bias which affects sample estimates so that they differ from the true. Problems of technique in taking the samples may be considered as the second general classification.

TABLE 6. SAMPLE YIELDS TOGETHER WITH ESTIMATES PREPARED BY THE A. M. S. FOR FOUR CROP-REPORTING DISTRICTS IN NORTH DAKOTA, 1938.

A. Durum wheat

District	County	Number samples	Sample yield	A.M.S. estimate	Weighted mean difference	Standard error
3	Cavalier	22	16.5	12.5		
	Grand Forks	12	21.1	19.5		
	Nelson	14	10.2	12.0		
	Pembina	10	16.8	18.6		
	Ramsey	18	12.4	11.5		
	Towner	26	16.0	13.0		
	Walsh	32	17.3	20.2		
	For district 3				.49	1.08
5	Eddy*					
	Foster					
	Kidder	6	11.7	6.0		
	Sheridan					
	Stutsman	4	11.6	6.2		
	Wells	2	12.7	6.5		
6	Barnes	2	14.7	13.0		
	Cass	2	15.1	17.0		
	Griggs					
	Steele					
	Traill					
9	Dickey					
	LaMoure	4	9.7	6.2		
	Logan	2	13.8	6.2		
	McIntosh					
	Ransom	4	12.0	11.5		
	Richland	4	15.1	14.6		
	Sargent	2	31.0	14.4		
	For districts 5, 6, and 9				4.19	1.39
For all districts					1.20	.84

* No samples were taken in the counties for which blanks are shown, i. e., no Durum samples in part A of table, and no other spring wheat in part B.

TABLE 6—Continued.

B. Other spring wheat

District	County	Number samples	Sample yield	A.M.S. estimate	Weighted mean difference	Standard error
3	Cavalier	6	18.7	12.5		
	Grand Forks	36	14.4	17.5		
	Nelson	4	5.5	11.0		
	Pembina	20	16.4	17.0		
	Ramsey	2	14.6	9.0		
	Towner					
	Walsh	22	13.3	16.6		
	For district 3				1.89	1.24
6	Barnes	34	14.9	8.6		
	Cass	52	21.1	13.5		
	Griggs	6	7.0	7.0		
	Steele	24	13.3	12.0		
	Traill	22	19.1	17.2		
	For district 6				4.94	1.41
5	Eddy					
	Foster					
	Kidder	4	14.3	4.6		
	Sheridan					
	Stutsman	20	11.1	4.6		
	Wells	2	6.2	5.2		
9	Dickey					
	LaMoure	4	15.2	4.0		
	Logan	14	7.3	4.0		
	McIntosh					
	Ransom	4	14.1	8.0		
	Richland	10	15.0	12.0		
	Sargent	2	29.2	8.8		
	For district 5 and 9				5.95	1.40
For all districts					3.02	1.06

BIAS

An idea of the extent of the bias in the sampling may be gained from the foregoing table which gives the yield estimates regularly issued by the Agricultural Marketing Service and the averages prepared from the objective sample.

The estimates prepared by the Agricultural Marketing Service divide the wheat into two types, Durum and other spring wheat. The differences, weighted by the number of samples in a county, between the two estimates were computed. The weighted mean difference and its standard error as shown in table 6 were then determined. These statistics were computed separately for the districts in which the most samples were taken. The remaining districts in which the sampling was light were also combined for this computation as shown in the table.

Obviously, it is impossible to obtain an exact measure of the amount of bias in the sample unless the true county and district yields are known so that the estimates of these yields determined by the sampling may be compared with the true values. In this investigation the actual yields per acre of the geographic units are not known. The Department of Agriculture obtains indications of the yield per acre by sending inquiries through the mail to farmers. The returns from this crop correspondence, adjusted by census returns, give an independent estimate of the yield which is generally believed to be fairly accurate on a statewide basis. A comparison of this estimate with the objective sample as shown in table 6 indicates that, in general, the yield estimated from the objective sample is higher. For the Durum wheat as a whole this difference is not significant. The difference is significant for the other spring wheat as a whole. Sub-groupings of the counties by districts indicate where these sampling differences occurred geographically.

While the amount of the bias cannot be exactly measured, because the actual yields are not known for the area sampled, some of the sources of the bias may be indicated. Bias due to the observer may be an important element. Throughout the sampling every precaution was taken to prevent the use of personal judgment. Yet, an analysis shows a significant difference between the samples taken by the two samplers.

The sampling was started with the assumption that the Crop Reporting Board's estimate of harvested acres excluded all fields and areas within fields that did not produce grain. Therefore, the samplers proceeded to exclude all bare spots within fields from the sample. During the field sampling it became apparent that, even though the schedules distributed by the Department of Agriculture called for harvested acres, there were, no doubt, some cases where the farmers reported the bare spots within the fields as harvested acres. In one case, the farmer stated that since the machine was run over the entire field, he considered the area in the field as harvested, even though one-third of the field in his case did not produce grain. In the south central part of the state, where the fields were thin and the yield light, a considerable amount of judgment was used in determining whether or not to include some fields and parts of fields in the sample. The judgment of the samplers in some of these cases may not have been the same as that of the farmers who reported to the Department.

It was evident from observation that yields adjacent to the roads were lower than yields occurring farther back in the fields. This was especially noticeable in the areas where there was a heavy infestation of grasshoppers. It appeared that grasshoppers were doing more damage around the border of a field than in the center. In taking the field sample, the first 20 paces from the road were excluded. This border effect raises the question—Can the population be limited to a strip lying parallel to the highway? A separate sampling study should be undertaken to determine the extent of the bias resulting from such a method of sampling the fields.

Another source of bias may be due to the expansion from the sampling unit. Magnification of errors by a factor, such as 10,000, may introduce a bias of considerable magnitude in the absolute sense into the sampling. The use of such a small sampling unit for field sampling may not be desirable. Present evidence concerning it is based only on the results obtained from sampling experimental plots. These plots are rather more uniform and homogeneous in their soil composition than farm fields.

TECHNIQUES

One of the first problems in technique is the securing of mature samples by the route sampling method just prior to, or at, harvest time. It was found in 1938 in the eastern half of North Dakota that within a single county the fields did not differ more than 7 days in date of maturity. About two-thirds of the fields did not vary more than 4 days in maturity. The greatest deviation in date of maturity within a county was due largely to varietal differences. The Durums on the average were about 3 days later than the bread wheats. There was a marked gradation in the date of harvest from the southern to the northern part of the state. In some areas the fields (especially the fields of Durum) were cut when the grain was in the dough stage in order to avoid grasshopper damage. In these areas it was decided to sample fields that would otherwise have been eliminated from the sample because of immaturity. Ordinarily if an immature field was selected for sampling, it was discarded, and a sample was taken from the nearest mature field along the route. This taking of samples from mature fields only at the time of sampling might be considered a possible source of bias. However, immature fields were selected no more than once or twice per 100 fields sampled. Thus, the substitution of mature fields for these immature selections would make the bias from this source very slight.

A few of the fields selected for sampling were already cut and shocked or windrowed. This did not present a difficult problem, as the location of the sampling unit could be made in the same way as the sampling unit would be located in standing grain. The number of heads was then determined by a "stubble count" of the sampling unit. The same number of heads was chosen from the nearby bundle or windrow (a deduction had to be made for the number of heads clipped by grasshoppers that lay on the ground in the sampling unit). Although this method of sampling was followed in 1938 and proved practical, it was much more time-consuming than the sampling of the uncut fields. If the field selected for sampling was already harvested by the combine, the next nearest field along the route was sampled.

Route sampling proved itself practical in 1938. True, it does not give a strictly random selection of fields, but completely random choice has the practical objections of time and cost involved in securing it. Furthermore, route sampling has several desirable features. It permits keeping the miles driven in each area proportional to the area of the geographical unit. Crop metering, at the same time, controls the sampling by keeping the number of samples taken proportional to the area in wheat. Such sampling then gives an indication of the distribution of the varieties over the area sampled. Route sampling also permits traveling over the area being sampled with the gradient of ripening. With increased experience in sampling it may be found advisable to adjust the sampling to the variability in an area. That is, an area with great variability would be sampled more intensively than a uniform region having the same area of wheat.

The success of estimating yield per acre of wheat by an objective method of sampling depends in part on how near harvest time the sample is taken. If the sample is cut and removed from the field before the grain has completely filled, there is a possibility that the mean yield of the sample will be below that of all farms as a whole because of the difference in the weight of the grain. Quality is also no doubt affected by early cutting. However, it is the opinion of cereal chemists that the elements which ultimately constitute the grain are largely translocated to the grain some time before the normal harvest. The yield or quality is therefore not likely to be greatly affected by pre-harvest sampling provided the heads are cut within 5 days of the normal harvest. In fact, several commercial companies in the wheat trade are now using such a method to determine the quality of wheat. During the sampling in 1938, in practically every instance, the samples were taken within 5 days of harvest. Upon the experience acquired in 1938 it appears that route sampling will be satisfactory and make it

possible to obtain sufficiently mature samples at the desired time. It may be well to point out that it has not been definitely proved just how much the quality and production are affected by harvesting at different stages of growth.

It is probable that the field samples were somewhat biased upwards because of the difference in the amount of wheat lost during threshing. There was practically no loss in threshing the grain of the sample which is in contrast to the amount of loss that normally occurs in the threshing and handling of grain on the farms. It is not known exactly how much grain is lost by harvesting the crop. Considerable experience will need to be acquired before the sampling can be adjusted for a bias coming from this source.

The size and shape of sampling unit within the fields presents another problem in technique. Yates and Zacopanay (3) found in testing different sizes and shapes of sampling units that a unit of one-half meter by four rows gave the maximum efficiency. The sampling unit used in this study, a U-shaped bar, 24" x 26.14", or approximately 1/10,000 of an acre in size, was convenient to handle in the field. The shape was such that 26.14 inches of four adjacent drill rows made up the sample, thereby including in the sampling unit the variability between rows. A "rod row," or a single drill row 1 rod long, gives about 20 percent more drill row than the rectangular unit used in this project. But the single drill row would not sample the differences due to competition among rows. On the other hand, the rod row sample might include greater variability due to soil heterogeneity. A sampling unit which is as representative as possible of the whole field will clearly give a better estimate than one which is representative of only a small part of the field. Hence, other things being equal, it would be desirable to ensure that the sampling unit include a maximum range of conditions existing in the field. The data used in the study of Yates and Zacopanay, mentioned above, pertained only to experimental plots that did not show any evidence of fertility gradient. Consequently, this shape of unit should be tried out to see if it is the most efficient under field conditions. Investigation to determine the comparative efficiency of sampling units of different sizes and shapes will add much to the available experience in sampling.

Although agronomists in sampling their experimental plots have not found a bias resulting from the use of a sampling unit as small as 1/10,000 of an acre, such a unit should be thoroughly tested under commercial conditions. Testing should tell if it is possible to make measurements accurate enough to permit a conversion to an absolute per acre basis without a systematic error entering into values determined from the samples. The components of this error would be the combined effects of the expansion from the sampling unit and the differences in losses in harvesting. Testing the accuracy of this small sampling unit by choosing a number of them at random in fields where the production is accurately determined should be worthwhile. A comparison between the sample mean and the actual mean would then give a basis for estimating the amount of bias resulting from using this unit.

In concluding the discussion of these problems in wheat sampling, it is pertinent to say that the accumulation of experience in the work over time will be the best guide to future methods. Ten years' data will make possible a far better evaluation of the bias and will point out the techniques which give the best results. As this experience is built up by the research section of the Agricultural Marketing Service, it may be incorporated into the regular procedures of the Service's work.

CONCLUSIONS

1. The investigation has shown that route sampling of the wheat crop to estimate and forecast yield per acre is a practical and an efficient method.

2. It was found that stratification by varieties would have resulted in a marked gain in accuracy. With stratification by varieties about 40 percent less fields would have been required to give the same precision. Geographical stratification would have added little to the information in the 1938 season.

3. The investigation showed the variance between fields to be larger than that within fields (mean squares of 65 and 19). The gain in accuracy would be small with increased sampling per field. Therefore, the sampling per field was adequate under the 1938 conditions. Sampling more fields with these conditions would add more to the information than increasing the number of samples within a field.

4. The regression analysis of the 1938 data showed number of heads per sample to be the best indicator of yield. The height of grain in the sample and the average length of heads added some information.

5. The yields determined from the objective sampling study exceeded very slightly the current estimates issued by the Department of Agriculture. Additional research is needed to determine the consistency and extent of this bias.

REFERENCES

(1) Quisenberry, K. S. Some plant characters determining yields in fields of winter and spring wheat in 1926. Jour. Amer. Soc. of Agron., 20:492-499. 1928.

(2) Yates, F. The place of quantitative measurements on plant growth in agricultural meteorology and crop forecasting. Report of the conference of Empire Meteorologists, London. 1935. Meteorological Office Publication No. 393.

(3) Yates, F. and Zacopanay, I. The estimation of the efficiency of sampling, with special reference to sampling for yield in cereal experiments. Jour. of Agr. Sci., 25:545-577. 1935.

(4) Yates, F. Crop estimation and forecasting: Indications of the sampling observations on wheat. Jour. Ministry of Agr., 43:156-162. 1936.

(5) Yates, F. Applications of the sampling technique to crop estimation and forecasting. Paper read before Manchester Statistical Society. Dec. 9, 1936.

(6) Cochran, W. G. Discussion "Crop estimation and its relation to agricultural meteorology." Supplement to Jour. Roy. Stat. Soc., 5: No. 1. 1938.

(7) Kalamkar, R. J. Jour. Agr. Sci. (1932), 22, 783.

(8) Kiesselbach, T. A. Studies concerning the elimination of experimental error in comparative crop tests. Nebr. Exp. Sta., Res. Bul. 13:1-95. 1918.

(9) Michels, C. A. and Schwenderman, John. Determining yields on experimental plots by the square yard method. Jour. of Amer. Soc. Agron., 26:993-1001. 1934.

(10) Sprague, H. B. Correlation and yield in bread wheats. Jour. Amer. Soc. Agron., 18:971-996. 1926.

(11) Hayes, H. K., Aamodt, O. S. and Stevenson, F. J. Correlation between yielding ability, reaction to certain diseases, and other characters of spring and winter wheats in rod row trials. Jour. Amer. Soc. of Agron., 19:896-910. 1927.

(12) Bridgford, R. O. Correlation of factors affecting the yield in hard red spring wheats. Master's thesis, Univ. of Minn. 1930.

(13) Bridgford, R. O. and Hayes, H. K. Correlation of factors affecting yield in hard red spring wheat. Jour. Amer. Soc. Agron., 23:106-117. 1931.

(14) Immer, F. R. and Ausemus, E. R. A statistical study of wheat and oat strains grown in rod row trials. Jour. Amer. Soc. Agron., 23:118-131. 1931.

(15) Laude, H. H. Relation of some plant characters to yield in winter wheat. Jour. Amer. Soc. of Agron., 30:610-615. 1938.