THE EXPERIMENTAL ERROR OF FIELD TRIALS.

BY W. B. MERCER, B.Sc. (*Vans Dunlop Scholar, University of Edinburgh*), AND A. D. HALL, M.A., F.R.S.

The Rothamsted Experimental Station.

THE magnitude of experimental error attaching to one or more field plots is a question of extreme importance in Agricultural Science, because upon its proper recognition depends the degree of confidence which may be attached to the results obtained in field work. A very cursory examination of the results of any set of field trials will serve to show that a pair of plots similarly treated may be expected to yield considerably different results, even when the soil appears to be uniform and the conditions under which the experiment is conducted are carefully designed to reduce errors in weighing and measurement.

From the Rothamsted plots one of us has already shown (*Journ. Bd. Agric.* 1909, XVI, 365) that the probable error attaching to a single plot is in the neighbourhood of plus or minus 10 per cent. That is to say, that the chances are even against the result being within 10 per cent. of the truth, by the truth being meant the average result obtained by a number of experiments large enough to eliminate casual variations.

There are two methods of reducing the experimental error to a negligible amount, one by repeating the experiment over a long period of years, the method which has been adopted with the old experimental plots of Rothamsted, or secondly, by multiplying the number of plots on a given field. The first method does not eliminate causal variations induced by type of soil, the second is also subject to errors due to the season. For example, in manurial trials, phosphates are known to give higher returns in wet than in dry seasons, so that however accurately the value of particular phosphatic manurings might be determined by a series of plots in one year only, that value is only true for the season in question, and if a more general or average

Journ. of Agric. Sci. IV   8

value is to be obtained, the experiments must be prolonged until the variations due to season have been eliminated. Again, in a comparison of varieties certain varieties may be particularly suited to a heavy type of soil; hence a trial in a single field, however exactly it expressed the relative values of the two varieties on that land, cannot be taken as holding for more than the one type of soil. Of course the type of soil may be regarded as a factor that must always be expressed in connection with any statement as to the value of a given manure or a given variety, and it is fundamentally incorrect to try and smooth this effect out by averaging results obtained on different types of soils. For example, a given variety equally adaptable to all soils may have its value truly represented by the figure 100, and another variety may have a value of 130 on a heavy soil and only 70 on a light soil. To express these latter results by an average value of 100 is to introduce an unnecessary error, obscuring the real character of the variety. Seasonal variations, since they depend upon an unpredictable factor, cannot however be thus eliminated in making any statement about the value of a given manure or a variety. We are driven to content ourselves with a result that is true over an average of seasons, because a farmer is obliged to assume beforehand that he will obtain such an average season. Granting, however, the necessity of repeating an experiment for several years if any result that is finally true is to be obtained, there are yet many occasions and many types of experiments in which it is desirable to obtain reasonably accurate conclusions from a single season's work, and the practical point at issue is the size and number of plots which are most likely to reduce the inevitable error within working limits. In order to obtain light on this question, during the year 1910 an attempt was made at Rothamsted to estimate the variations in the yield of various sized plots of ordinary field crops which had been subjected to no special treatment and appeared to the eye sensibly uniform. The fields were selected, one of wheat which promised to be a fair crop for the season and was generally standing up well, the other of mangolds which looked a uniform and fairly heavy crop for the season and soil. In the wheat field a very uniform area was selected, one acre of which was harvested in separate plots, each one five-hundredth of an acre in area. The small sheaves which each plot yielded were then stored and eventually threshed out by a hand machine, corn and straw being separately weighed.

In measuring the plot a fixed number of rows of the drill were taken, eleven in this case, and a fixed length of 10·82 feet measured

TABLE I. *Plan and yields in lbs. of mangold plots; upper figures roots, lower leaves.*

S.

376 / 67	371 / 52	355 / 50	356 / 50	335 / 48	332 / 44	335 / 51	340 / 50	352 / 46	352 / 49
316 / 56	338 / 51	336 / 48	356 / 50	332 / 49	340 / 48	339 / 47	359 / 48	365 / 52	349 / 46
326 / 56	326 / 47	335 / 47	343 / 49	330 / 47	328 / 47	350 / 52	339 / 47	337 / 50	362 / 46
317 / 54	343 / 54	330 / 50	327 / 46	336 / 45	340 / 51	335 / 47	301 / 44	348 / 46	357 / 52
321 / 50	332 / 46	317 / 44	318 / 47	306 / 48	331 / 46	335 / 47	311 / 47	329 / 46	353 / 49
335 / 56	349 / 51	332 / 44	358 / 48	325 / 45	348 / 49	328 / 48	290 / 41	326 / 44	323 / 44
341 / 57	316 / 52	321 / 46	321 / 48	314 / 45	340 / 50	337 / 52	321 / 45	329 / 44	317 / 41
347 / 59	363 / 58	349 / 49	341 / 51	329 / 46	321 / 46	337 / 45	300 / 46	364 / 50	310 / 45
354 / 64	365 / 54	350 / 49	339 / 47	316 / 42	342 / 49	335 / 49	331 / 48	345 / 51	327 / 47
309 / 55	352 / 56	328 / 49	297 / 43	308 / 40	343 / 46	349 / 54	372 / 51	362 / 55	346 / 51
349 / 60	335 / 50	323 / 48	325 / 43	322 / 46	348 / 54	350 / 50	354 / 51	350 / 52	360 / 51
343 / 57	346 / 52	300 / 41	312 / 48	313 / 42	331 / 51	347 / 47	324 / 46	336 / 46	339 / 44
338 / 53	327 / 49	323 / 45	296 / 38	311 / 40	335 / 47	320 / 44	318 / 44	343 / 44	333 / 45
339 / 52	384 / 61	322 / 46	331 / 44	277 / 37	299 / 43	332 / 50	302 / 39	318 / 42	306 / 41
314 / 52	325 / 46	330 / 48	328 / 44	301 / 40	316 / 43	338 / 46	310 / 45	304 / 42	302 / 37
300 / 52	329 / 50	322 / 48	310 / 45	309 / 40	342 / 49	350 / 51	335 / 46	320 / 42	278 / 38
369 / 60	341 / 51	344 / 47	324 / 44	316 / 44	342 / 51	351 / 53	324 / 46	310 / 47	309 / 44
298 / 49	302 / 47	325 / 47	300 / 46	286 / 40	330 / 52	324 / 49	322 / 43	309 / 45	322 / 49
324 / 52	316 / 48	310 / 45	317 / 44	295 / 40	308 / 45	295 / 54	267 / 40	307 / 43	290 / 41
330 / 57	306 / 48	306 / 45	306 / 41	291 / 41	321 / 45	325 / 48	288 / 41	302 / 46	310 / 42

E. (left)   **W.** (right)

N.

8—2

along the rows for cutting. Thus, if there were any variation in the breadth of the drills the assumed one five-hundredth of an acre would not represent the actual land area of each plot. These variations were small: moreover in all experimental crop work it is necessary to see that comparative plots do contain exactly the same number of drills, so as to make the measured area one of crop rather than of land.

For the second trial, mangolds, the area in the same way was divided up into plots one two-hundredth of an acre, leaves and roots being weighed separately. As before, each unit plot contained the same number (three) of drills and was measured along the drill a fixed length of 30·25 feet.
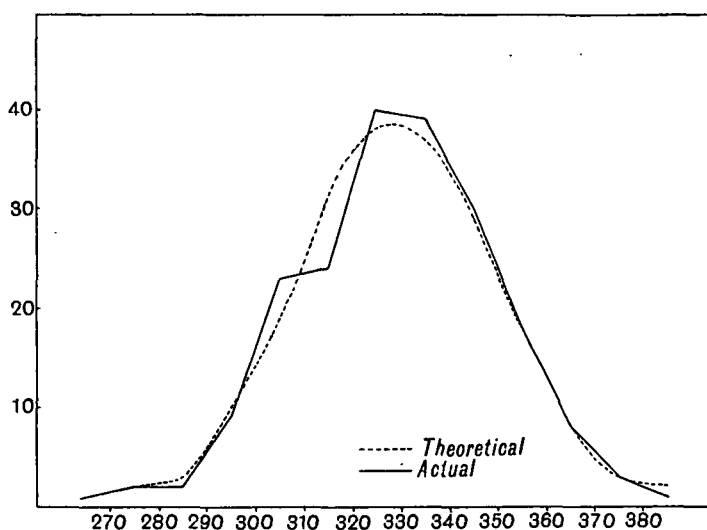
Fig. 1. Frequency curve for 200 plots of mangold roots; actual and theoretical.

It will be convenient to begin with mangold experiments which contain the smaller number of plots, but the principles discussed are the same in both cases:

Table I shows the weights of roots and leaves on all the plots as they stood in the field; the length of the plots runs with the horizontal lines of figures, this being also the direction of the drills across the field. A first examination of the results shows that large variations occur in the weights of individual plots; the highest yield of roots is 384 pounds and the lowest 267, variations of 18 per cent. on either side of the mean. In order to get a general idea of the magnitude and nature of the variations it is convenient to plot them into a frequency curve, and

Fig. 1 shows the curve obtained by gathering the weights of the roots into groups at intervals of 10 lbs., together with the "normal curve of error," calculated to fit the observations. It will be seen that the two curves agree very closely, considering the small number of observations, and there is no sign that the material is not homogeneous. This enables us to apply the well-known properties of the normal curve of error to our results. The next thing which may be examined is whether the position of the plot in the field has any effect upon its yield, *i.e.* whether there is any uniform change in the character of the soil, etc., that brings about a steady change in the weights in passing either down or across the field. This can be roughly tested by adding together the weights on each plot as set out in the table, first horizontally and then vertically.

The results are set out in Table II, and expressed graphically by

TABLE II.   *Varying weight of rows of plots; mangolds.*

Total Weight of Rows in lbs.	
East to West	South to North
6646	3504
6796	3430
6558	3376
6505	3334
6252	3253
6637	3314
6712	3287
6408	3361
6656	3404
6545	3366
	3416
	3291
	3244
	3210
	3168
	3195
	3330
	3118
	3029
	3085

the two curves in Fig. 2. It is clear that there is a slight falling off in yield as one passes from South to North, but the variations from East to West are covered by the experimental error. Where such a uniform change can be traced as shown from South to North in these experiments

it might be possible to apply a correction for position. But the device is cumbrous and the same end may be gained by duplication and reflection of the order of the plots about their centre. Thus duplicates of three plots would be arranged in a row in the order a, b, c, c, b, a. If the variation in the land is perfectly uniform the error due to position obviously disappears.

Further it must not be supposed that the results obtained for this set of plots in 1910 can be used to correct results obtained on the same plots in any other year. For example if any given area gives a yield of 10 per cent. over the mean in 1910 we cannot assume that we shall obtain a true average yield by deducting 10 per cent. from the yield of
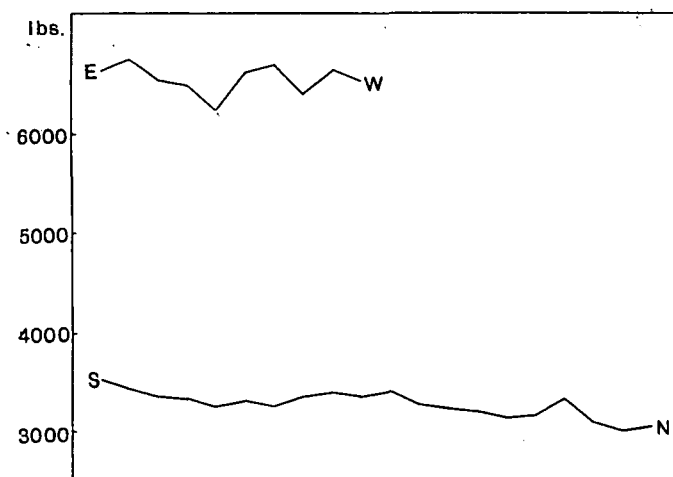

Fig. 2. Varying weight of rows of plots with position in the field. Mangolds.

the same plot in 1911. The variations, as we have learnt from other observations, are determined by so many incalculable factors that they may fairly be described as casual, and a plot above the average in one year is as likely as not to be below the average in the next year.

We may now test whether the weights of roots and leaves vary together. As the whole area had been similarly treated it might be expected that the variations in the weight of the leaves would follow fairly closely those of the roots. Moreover, if this fact can be established attention may be confined to the roots alone, as the same arguments will apply to both equally.

The plots were therefore gathered together into groups according to their weights, each group containing all the plots whose weights fell

within a range of 10 lbs., *e.g.* 379—370, 369—360, 359—350, etc. The average weight of the roots of the plots in each of these groups was then found, this of course being somewhere about five in the unit, and also the average weight of the leaves that accompanied the roots on each of the plots in question. Thus we obtain the average weight of leaves corresponding to the average weight of roots for a series of groups of plots giving regular increases in their yield of roots. These results are set out in Table III and are plotted graphically in Fig. 3, forming

TABLE III. *Relation between Roots and Leaves.*

No. of Plots in Group	Mean Weight of Group in lbs.		No. of Plots in Group	Mean Weight of Group in lbs.	
	Roots	Leaves		Roots	Leaves
1	267	40·0	39	334	48·3
2	277	37·5	30	344	50·8
2	287	40·5	18	353	51·1
9	295	43·3	8	364	53·3
23	304	44·1	3	373	56·6
24	314	45·5	1	384	61·0
40	324	46·6			

what is known as a Regression Line. If the weight of leaves had been exactly proportional to the weight of roots we should obtain a straight line on the graph and it will be seen that the experimental material does yield approximately a straight line, the deviations being only considerable at the extremes of the graph, when the numbers of plots making up the group are smaller and the casual variations are not smoothed out.

The character of the graph shows us that the weights of leaves and roots vary together, so that further discussion may be confined to the roots.

The next step is to determine the standard deviation of the plots from the mean, the standard deviation being the square root of the sum of the squares of all the differences from the mean after dividing by one less than the total number of plots. Two-thirds of the standard deviation gives us the probable error of a single plot, that is to say, the limits within which it is an even chance a particular plot will fall. If a, b, c, etc., be the weights of each plot, m their mean, and n the number of plots, then

$$S. D. = \sqrt{\frac{(a-m)^2 + (b-m)^2 + (c-m)^2, \text{ etc.}}{n-1}}.$$

Probable error = S. D. × ·6745 or 2/3 S. D. approximately.

Working in this way the probable error attaching to a single plot of one two-hundredth of an acre in area is found to be 20·4 lbs., the mean being 328·6 lbs., or as previously explained, it is an even chance that any given plot weighs more than 308·2 and less than 349 lbs.
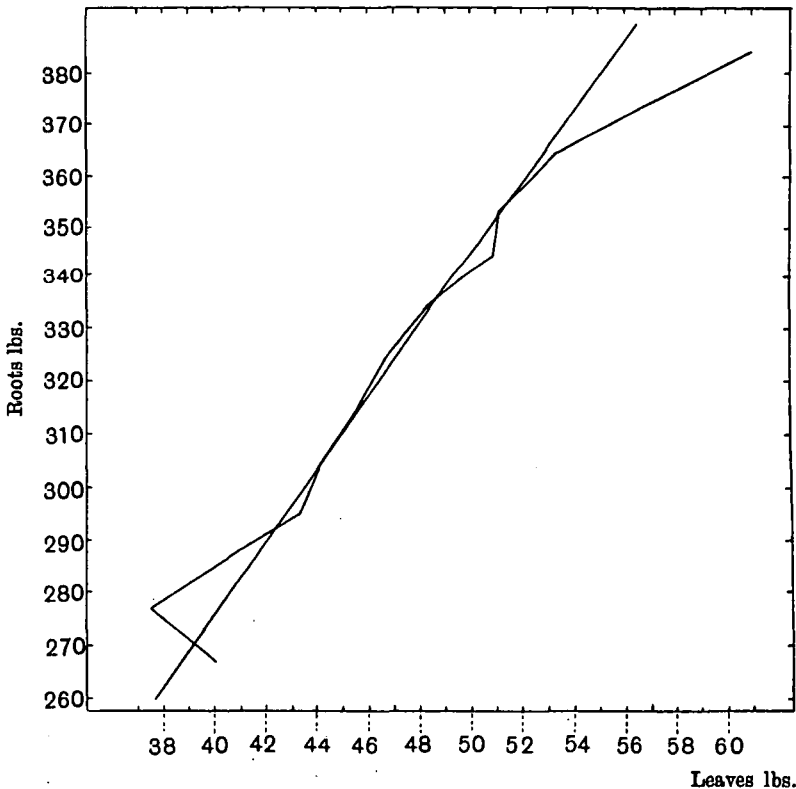

Fig. 3.   Regression line showing average weight of leaves for any weight of roots.

Calculating as a percentage of the mean the standard deviation becomes 6·2 per cent.   The next step is to group pairs of adjacent plots together to form single plots of twice the former length and one-hundredth of an acre in area.   When this is done the standard deviation is reduced to 5·36 per cent.   Again, grouping adjacent plots in pairs so as to obtain five plots in all from East to West, and 10 plots from North to South, thus making up areas of one-fiftieth of an acre, the standard deviation becomes 4·51 per cent. of the mean.   Grouping once more five adjacent plots, so as to obtain four larger plots from North to South and 10 from East to West, each one-fortieth of an acre

in area, the standard deviation becomes 4·06 per cent.; again grouping 10 adjacent plots together so as to get units of one-twentieth of the acre, the standard deviation falls to 3·6 per cent. of the mean, and is further reduced to 3 per cent. when the plots are made up to one-tenth of an acre[1]. It is clear that the probable error attaching to the result obtained for a single plot diminishes with the size of the plot. We can obtain some rather useful information by plotting the standard deviation in percentages against the size of the plot, which gives us
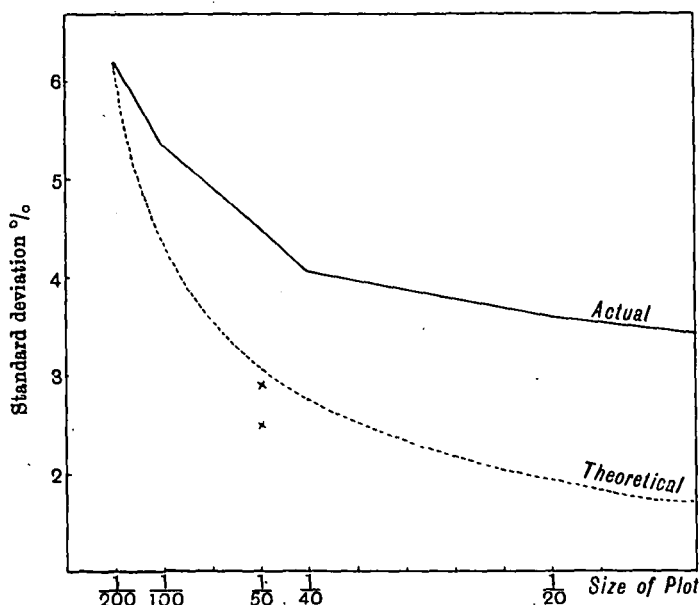

Fig. 4.   Percentage standard deviations with plots of different size ; mangolds.

the curve shown in Fig. 4. On the same figure is also drawn a theoretical curve that has been deduced on purely statistical grounds to give the standard deviation of successively larger plots made up of a number of the original one two-hundredths taken at random, *i.e.* not consecutively, nor in any other regular manner. The curve for the experimental results shows that the reduction of the probable error is not very rapid when a plot increases above a certain size. In fact the curve indicates that at about one-fortieth of an acre the

[1] As the number of plots decreases the value which can be attached to the determination of the S. D. diminishes, and in the later calculations when small numbers of plots are made up from scattered units, the results given are invariably averages of two or three determinations.

probable error attaching to a single plot has attained such a figure that it is not profitable to try and reduce it further by merely increasing the size of the plot.

As regards our standard plot then, one in the neighbourhood of one-fortieth of an acre is large enough.

We may next vary our groupings of the small one two-hundredth of an acre plots, so as to ascertain if the shape of the plot has any effect upon the result. This was done by making up plots of one-twentieth of an acre in two ways, one two plots long from East to West, and five plots deep from South to North (making a plot about 20 yards long along the drills and 12 yards broad), the other a long narrow plot, five of the small plots long from East to West, and two only deep from North to South, *i.e.* about 50 yards × 5 yards. For the first description of plot the standard deviation is 3·47 per cent., for the second 3·73 per cent., results from which little can be deduced as to any superiority of long and narrow plots over square ones.

Having thus ascertained the probable error attaching to plots of a particular size and shape it is now desirable to get some light upon the reduction in the probable error that can be effected by multiplying the number of plots. Multiplying the number of plots would be only equivalent to increasing their size if the ground were absolutely uniform, but as we have every reason to expect that the variations in the yield of individual plots are largely due to inequalities of the soil, these may be expected to be considerably smoothed out if the area making up the larger plot is sub-divided into a number of smaller plots scattered more or less regularly over the experimental area. The fact that the theoretical curve in Fig. 4 so quickly falls below the observed one, is evidence that some considerable correlation exists between the yield of adjacent plots.

We began by making up plots of one-fiftieth of an acre by grouping together the yields of four of the one two-hundredth of an acre plots scattered systematically. The first method of scattering adopted gave a standard deviation of 2·5 per cent. for the one-fiftieth of an acre plots; a second method of scattering in which the distribution was not quite so regular gave a standard deviation of 2·9 per cent. Comparing these with the previous results we see that the standard deviation attached to a plot of one-fiftieth of an acre has been reduced from 4·5 per cent. to 2·7 (mean) by taking a fiftieth of an acre in four scattered two-hundredths instead of in one block.

Thus a very fair measure of precision can be attained with as small

an area as one-fiftieth of an acre when it can be sub-divided into scattered plots. However, one two-hundredth of an acre is a very small plot for most kinds of field experiments, and we have already decided that the best size of plot comes somewhere in the neighbourhood of one-fortieth of an acre. Accordingly, a further calculation was made of the standard deviation of plots of one-tenth of an acre made up of four scattered fortieths. Two methods of scattering were adopted; one gave a standard deviation of 2·0 per cent., the other of 1·3 per cent.

Taking the mean of these figures [S. D. = 1·65 %], we find that tenth acre plots made up of four or five units scattered about the experimental field ought to yield us a result with a probable error of about one per cent. In other words, if the yield of this field is 30 tons per acre, then the weight obtained from five portions each one-fortieth of an acre taken at various places about the field, might be expected to fall between 30 tons 6 cwt. and 29 tons 14 cwt., the chances being even for and against such a measure of accuracy.

Again, as with simple plots the standard deviation for a plot made up of a number of smaller scattered units is not diminished in proportion to their increase in size, though for any given size of plot it is considerably reduced if the plot be made up of several scattered units. It is desirable to find whether the standard deviation will be decreased proportionately if the number of units into which an area is divided is increased. Determinations were therefore made of the standard deviation for plots of one-fifth of an acre made up of two-tenths, four-twentieths and ten-fiftieths respectively. The results are set out in Table IV, and expressed graphically in Fig. 5, standard deviation being plotted against number of units.

TABLE IV. *Standard deviation with different numbers of component units.*

No. of units	Standard Deviation
1	2·48
2	1·76
4	1·36
10	1·19

The figure shows that little is gained by increasing the number of scattered units beyond four or five, so that the results for one-tenth of an acre given above yield the lowest standard deviation which it is

economically possible to obtain. This is quite sufficient for ordinary experimental purposes, though, as we shall show later, in certain special cases where a very close comparison is necessary between a pair of varieties or two methods of treatment, a still greater precision can be obtained by a proper arrangement of the experimental areas.

It should be noted that the values obtained for the standard deviation for a given size and number of plots must not be interpreted too literally. They are only strictly applicable to the particular field, crop and season under discussion, because under other conditions the
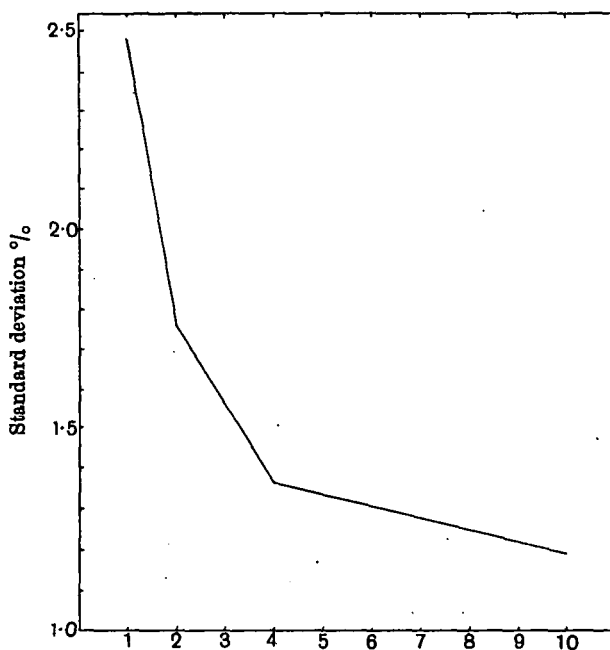

Fig. 5. Variation of standard deviation with number of units in plot ; mangolds.

factors inducing variation may become considerably altered. The principles, however, upon which they are founded hold in all cases, and the actual values obtained indicate in all probability the order of magnitude of the error that will be experienced.

We may now turn to the results obtained with wheat, where we have a larger number of plots, 500 in all. Table V shows the actual yields of grain and straw per plot in pounds, arranged as they occurred in the field. The maximum yield of grain is 5·16 lbs., the minimum 2·73 lbs., a variation of 30 per cent. on either side of the mean.

TABLE V. Plan and weights in lbs. of 500 wheat plots; upper figures grain, lower straw.

E. (top) **N.** (left) **S.** (right) **W.** (bottom)

Each cell shows upper figure = grain, lower figure = straw (grain/straw).

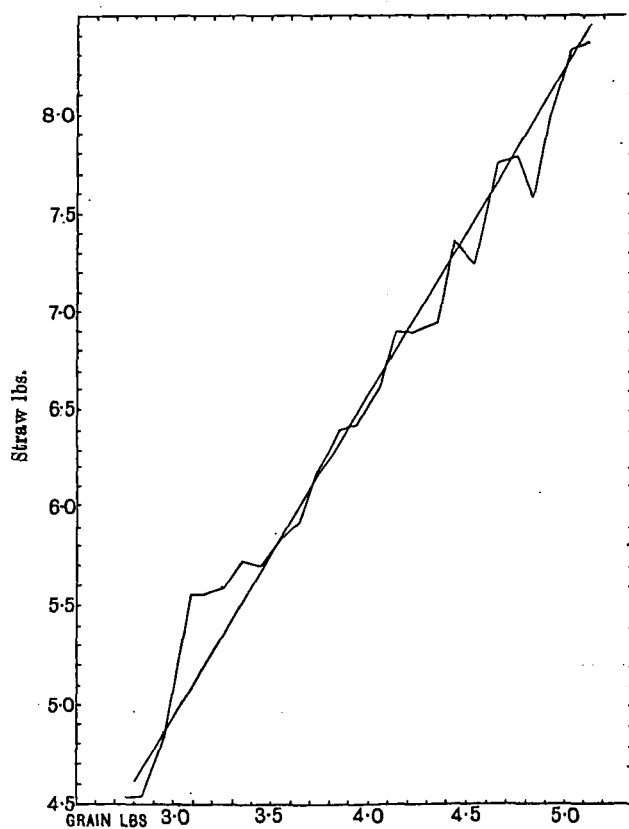| 4·53/6·78	4·23/6·08	2·97/4·53	3·17/4·95	3·36/4·76	3·72/5·47	4·27/6·54	3·51/5·05	3·57/5·12	3·66/5·96	3·64/5·11	3·92/6·33	4·58/7·23	4·02/6·10	4·37/6·75	4·00/5·87	4·14/6·98	4·04/6·08	4·75/7·91	4·48/6·08	3·04/4·71	5·13/7·99	4·06/7·19	4·15/6·85	3·63/6·37
3·94/5·68	4·09/5·91	3·14/5·11	3·53/5·09	3·69/5·56	4·36/6·14	4·26/8·61	3·92/5·83	3·44/5·00	3·82/5·80	3·61/5·58	3·97/6·03	4·05/6·57	4·19/6·43	3·97/6·09	4·50/6·50	4·03/6·91	4·27/6·35	4·56/7·88	3·74/6·63	4·03/6·34	4·64/7·80	4·15/7·41	4·21/7·29	4·07/6·24
4·38/7·49	3·97/6·28	3·59/5·41	3·66/6·15	4·33/6·17	4·31/6·44	4·20/6·55	3·53/4·91	3·44/5·56	4·07/6·74	3·64/5·86	4·58/7·23	3·73/6·02	4·07/6·37	3·94/6·18	3·67/6·20	3·30/5·95	3·76/5·93	4·76/8·18	4·46/6·98	3·77/6·17	4·69/7·50	4·40/7·35	4·29/7·71	4·51/7·05
3·64/5·55	3·46/5·41	3·32/4·62	3·89/5·92	3·81/6·13	4·10/5·96	4·30/5·95	3·77/5·79	3·97/6·03	4·54/7·08	3·75/4·62	3·19/6·56	4·06/6·19	3·34/5·60	4·01/5·99	4·69/5·41	3·05/5·88	4·52/7·29	4·52/7·60	3·91/6·46	3·49/5·70	4·04/6·66	4·05/7·89	4·64/8·23	3·90/6·91
4·31/6·56	3·96/5·60	3·57/5·24	3·67/5·45	3·55/5·89	3·74/5·76	4·22/6·15	4·29/5·58	3·79/5·33	3·92/6·14	3·70/7·67	4·14/5·98	3·74/6·13	3·63/6·43	3·83/6·36	5·07/8·05	3·67/7·33	4·17/7·33	4·40/7·91	4·10/6·77	3·76/6·05	4·64/7·86	4·92/8·58	4·27/7·73	3·63/5·93
4·29/6·15	3·75/5·50	4·24/5·82	3·94/5·87	4·20/5·92	3·58/5·61	4·24/5·88	4·35/6·09	3·87/5·57	4·01/5·99	3·37/5·00	3·76/5·49	3·72/6·03	4·09/6·10	3·79/5·46	4·36/5·58	4·26/7·61	3·84/6·28	4·39/7·36	3·66/6·15	3·61/5·89	4·73/7·98	4·08/7·04	3·55/6·45	3·16/5·59
3·44/5·68	2·78/4·28	3·44/4·56	3·05/4·57	3·42/5·45	4·06/6·25	3·81/5·69	4·38/6·24	3·97/6·03	4·59/7·28	3·71/5·66	3·77/5·48	4·33/6·79	3·74/6·38	3·96/6·23	4·86/7·51	4·82/7·37	4·06/6·81	4·94/8·06	3·56/6·06	3·28/4·97	4·39/6·98	4·23/7·02	3·50/5·87	3·18/5·32
4·24/7·26	3·47/5·59	3·76/5·24	3·84/5·66	3·95/5·92	3·66/5·78	3·54/5·21	3·71/5·91	3·83/6·29	4·76/6·49	3·71/5·79	3·93/6·07	3·72/5·97	3·04/4·96	3·93/6·13	4·93/7·57	4·07/6·99	3·40/6·10	4·44/7·50	3·84/6·10	3·72/5·78	4·66/7·28	4·07/7·05	3·35/5·71	3·42/5·52
4·55/6·32	3·44/4·81	4·17/5·58	3·63/5·81	4·21/6·04	4·41/6·15	3·85/6·15	4·21/6·29	4·24/6·45	4·73/8·64	3·75/5·12	3·96/6·35	4·05/6·82	3·59/6·03	3·56/5·75	5·02/8·23	4·37/7·25	2·99/5·07	4·64/7·92	4·11/6·83	3·75/5·94	4·49/6·95	4·67/7·64	3·61/6·01	3·97/6·03
3·91/5·96	2·84/4·10	3·97/5·65	4·03/5·72	4·08/7·00	3·95/6·11	4·09/5·47	4·07/6·18	4·09/6·16	4·20/6·49	3·09/5·47	3·47/5·78	3·37/6·25	3·99/6·26	3·86/6·14	4·44/6·75	3·99/7·26	3·61/6·83	4·66/7·59	4·20/6·86	4·13/7·31	4·42/6·95	4·27/7·17	3·71/6·29	3·40/5·66
3·29/5·65	3·47/5·84	3·89/5·98	4·08/6·54	4·24/7·20	4·54/7·46	4·36/7·39	3·86/6·14	3·41/5·78	3·74/8·63	3·37/6·44	3·29/5·71	4·09/7·28	4·17/7·08	4·17/7·20	4·39/7·73	4·17/7·52	3·95/7·11	4·77/8·23	4·21/8·23	4·01/7·18	4·51/7·86	3·84/6·60	3·64/6·30	3·39/5·61
3·83/5·99	3·56/6·38	4·07/5·80	3·97/6·47	4·11/6·95	4·47/7·84	3·94/6·68	3·47/5·90	3·51/6·11	4·36/7·26	4·86/6·39	3·14/5·05	3·99/7·13	3·07/6·12	4·10/6·90	4·56/7·73	3·72/7·28	4·08/7·17	4·45/8·74	4·37/8·75	3·59/6·53	4·45/7·17	3·82/6·87	3·89/7·05	4·23/7·14
4·33/7·11	3·93/5·69	4·02/6·35	4·58/6·79	4·11/6·64	4·54/7·33	3·67/6·20	3·89/6·23	4·13/6·87	3·60/6·27	3·39/5·86	3·05/5·70	4·09/7·72	3·91/7·34	3·99/7·26	4·63/7·87	3·86/7·20	3·92/6·70	4·42/8·02	4·47/8·53	3·94/7·06	4·57/7·93	4·41/7·03	4·09/7·22	4·52/7·10
3·93/6·07	3·47/5·78	4·39/6·11	3·91/6·09	3·35/6·27	3·65/6·60	3·05/5·20	3·80/6·14	3·19/5·87	3·60/5·84	3·64/6·36	3·43/5·38	3·56/6·69	3·79/6·33	4·36/7·39	4·47/7·16	3·73/6·89	3·44/6·62	4·37/7·69	3·96/7·10	4·29/7·06	4·08/6·67	4·39/7·73	4·42/5·20	4·46/7·16
3·38/5·68	3·43/5·82	3·69/5·18	3·25/6·37	3·71/6·29	3·86/6·64	3·48/5·52	3·48/5·52	3·05/5·85	3·39/6·42	3·30/5·70	3·47/5·84	3·41/5·96	3·55/6·70	4·27/7·17	3·96/7·10	4·13/7·24	4·11/7·58	3·89/6·86	3·54/6·46	3·96/6·54	4·32/6·93	4·29/7·52	3·28/7·03	5·13/8·37
3·63/5·99	3·79/6·21	3·80/5·70	3·69/5·93	3·22/5·40	3·91/6·71	3·26/5·25	3·23/5·33	2·92/4·95	3·45/6·05	3·51/5·99	3·82/6·37	3·16/5·84	4·06/7·00	3·41/5·96	4·47/7·84	4·13/7·24	4·12/7·82	4·27/7·79	4·29/6·96	3·78/6·72	4·58/7·73	3·89/7·05	3·87/6·82	4·23/6·89
3·83/5·99	3·68/6·01	4·21/5·48	3·87/7·50	3·87/6·13	3·77/6·98	3·05/5·85	3·68/5·82	3·05/7·64	3·14/5·48	3·60/6·34	3·91/6·21	3·75/6·31	4·19/7·30	4·52/7·73	4·61/7·51	4·21/7·41	4·41/7·53	3·91/7·34	3·49/6·63	4·84/8·85	3·92/6·70	4·23/7·14	4·12/7·38	4·38/6·72
3·67/6·33	3·84/6·35	3·69/5·43	3·76/6·11	3·48/6·14	3·36/6·08	3·66/5·84	3·09/5·41	2·73/4·77	3·01/5·68	3·54/5·58	4·41/6·78	4·49/7·57	5·13/8·31	4·68/8·07	4·91/7·96	4·15/7·47	4·37/7·19	4·68/7·51	3·94/6·75	4·46/7·54	5·16/8·78	4·69/8·06	4·28/7·03	3·85/6·59
4·19/6·93	4·21/6·98	4·33/6·04	3·91/6·09	3·93/6·57	3·71/6·35	3·68/6·76	3·36/6·14	2·85/4·96	3·29/5·58	3·24/5·95	4·28/6·97	4·19/6·93	4·52/8·17	4·44/6·87	4·46/6·91	3·89/7·36	3·97/7·53	4·06/7·06	4·41/8·15	3·66/7·09	5·09/8·72	4·42/8·45	4·22/7·65	3·61/6·20
3·36/6·33	3·76/6·36	3·59/4·66	3·54/6·21	3·71/6·35	3·36/6·08	3·49/5·82	3·49/5·82	3·48/5·52	3·14/5·48	3·24/5·95	4·28/6·97	3·70/6·80	4·52/8·17	4·44/6·87	4·46/6·91	3·89/7·36	3·97/7·53	4·06/7·06	4·22/7·72	3·66/7·09	5·09/8·72	4·42/8·45	4·22/7·65	3·61/6·20

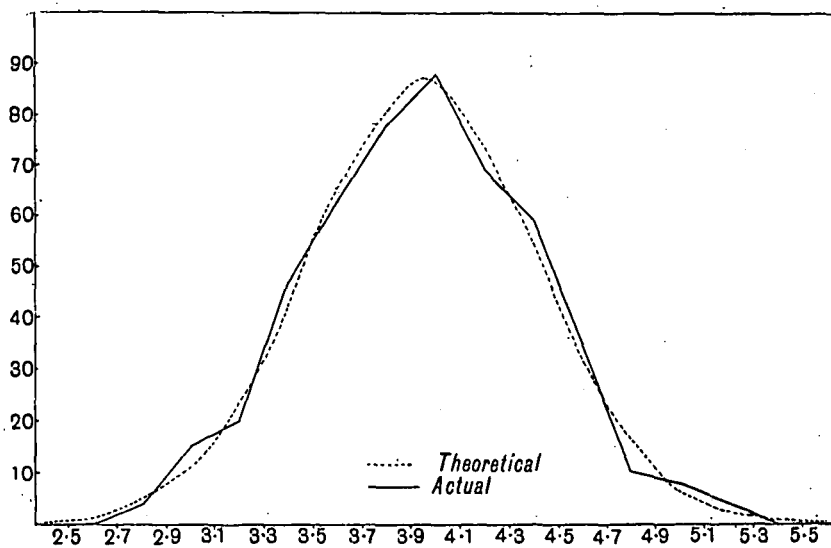Fig. 6.   Regression line showing average weight of straw at any weight of grain.

Fig. 7.   Frequency curve; wheat.

Table VI and Fig. 6 show the relation of grain to straw, obtained by the method previously described, the plots having been grouped according to the weights of grain, at intervals of 0·1 lbs., *e.g.* 2·7—2·79, 2·8—2·89, 2·9—2·99, etc., and the average weights of grain and straw on the plots falling in each group being determined. Again a somewhat irregular regression line is obtained, though it lies sufficiently close to the straight line to show that the yields of grain and straw do vary together, and that there is no tendency to such an inverse variation as would cause the straw to be light where the grain is heavy. Hence it will be sufficient to deal with the grain alone.

TABLE VI.   *Relation between grain and straw.*

No. of Plots in Group	Mean Weight of Group in lbs.		No. of Plots in Group	Mean Weight of Group in lbs.	
	Grain	Straw		Grain	Straw
2	2·75	4·53	42	4·06	6·61
2	2·85	4·53	28	4·14	6·90
3	2·96	4·85	41	4·24	6·90
12	3·08	5·55	31	4·36	6·95
10	3·17	5·56	28	4·44	7·36
10	3·26	5·59	22	4·55	7·25
18	3·36	5·73	13	4·66	7·76
29	3·45	5·70	7	4·76	7·79
26	3·55	5·83	3	4·85	7·58
37	3·65	5·92	5	4·94	8·00
44	3·74	6·17	3	5·06	8·33
34	3·85	6·40	4	5·14	8·36
46	3·94	6·42			

Following the previous method of treatment we may now begin by plotting the result into the frequency curve, Fig. 7. This curve fits the theoretical one as well as may be expected considering the number of observations, so we may conclude that the material is fairly homogeneous. Adding again the weights from the plots in rows from East to West, and from North to South, the results of which are expressed in Table VII and Fig. 8, we again see evidence of casual irregularity. There is evidently no systematic variation of soil from East to West, and though from North to South there seems to be a decline in yield it is too irregular to be made a basis of any corrections. Certain systematic correlations were also tested. The field was laid up in lands 15—17 yards broad. There were four (or sometimes only

TABLE VII.   *Varying weight of rows, wheat.*

Total Weight of Rows in lbs.	
North to South	West to East
98·9	78·3
100·2	79·3
101·1	85·6
97·9	90·8
100·7	76·4
98·8	81·1
96·8	89·5
97·0	79·3
102·0	78·8
97·8	90·6
98·9	82·3
99·6	77·9
102·9	77·9
98·4	74·8
92·7	72·2
93·6	77·3
97·5	70·6
97·9	75·1
104·7	77·5
96·5	80·1
	76·3
	75·2
	75·9
	73·1
	78·9
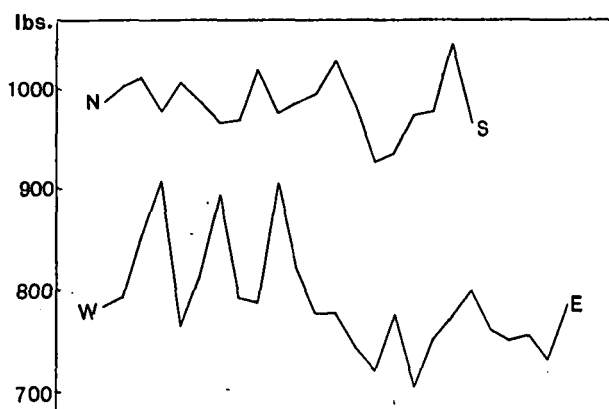
Fig. 8.   Varying weight of rows of plots with position in field ; wheat.

three) rows of plots on each land, a strip of varying breadth containing the actual furrow being left uncut between each row of plots. It was thus possible to obtain a comparison between the plots occupying the crown of the ridge and the neighbouring plots adjoining though not actually containing a furrow. Table VIII shows the mean result obtained in each case, from which it will be clear that the situation of the plots on the crown of the ridge or near the furrow had no effect upon their yield. This result is somewhat surprising, because to the eye the plots upon the ridge appeared to be decidedly better, an illusion we may assume to have been caused by the wheat ears on the ridge being lifted up to the eye.

TABLE VIII. *Influence of position on land.*

No. of Land	Mean Weight of Plots in lbs.	
	Middle of Land	Sides of Land
1	4·12	4·22
2	4·05	4·15
3	3·94	4·24
4	3·89	3·92
5	3·70	3·68
6	3·91	3·82
7	3·80	3·79
Mean	3·91	3·97

Again, one part of the area was a good deal contaminated with thistles, and notes were made at the time of cutting of the abundance of thistles or otherwise, in each plot. Altogether there were 187 plots in which the thistles were distinctly prominent, and these gave an average yield of 3·79 lbs. of grain, as against 4·04 lbs. of grain on 313 plots containing few or no thistles. The standard deviation of the difference may be calculated from the S.D. of the whole series of plots (0·46 lbs.) as follows: $0\cdot46 \sqrt{\frac{1}{187} + \frac{1}{313}}$ or ·041 lbs., so that the probable error of the difference is ·027 lbs. Thus the presence of thistles had depressed the yield by $0\cdot25 \pm 0\cdot027$ lbs., or 6·16 per cent., a result which we can accept with some confidence because it is 9 times greater than the probable error. This appears to be a very small reduction of yield for a very marked contamination of the plots

with thistles, but it confirms the opinion of many farmers that thistles do less injury to a corn crop than other weeds like couch.

Every operation in harvesting must result in additional error, and since grain and straw vary together, it seemed possible that the weights of total produce would give a better idea of the relative value of different plots, or that the standard deviation calculated for grain + straw would be smaller than that for grain alone. But the standard
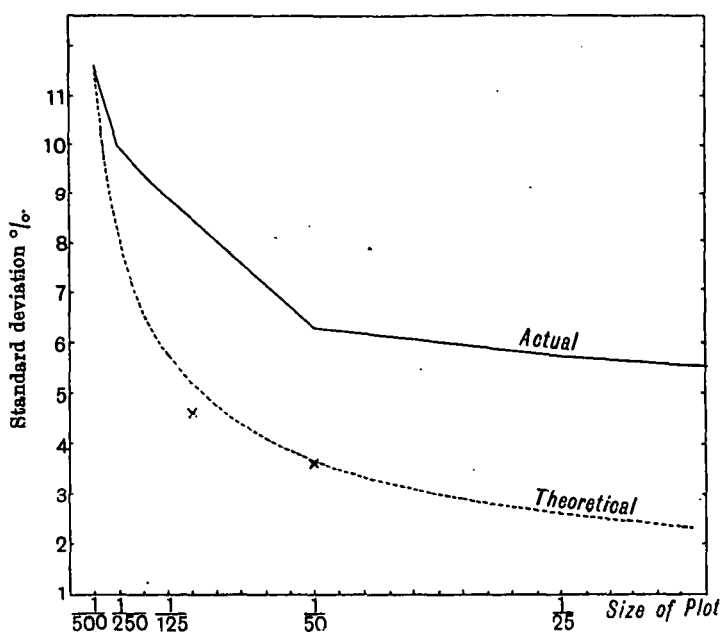

Fig. 9. Standard deviation in plots of different sizes; wheat.

deviation of the total produce proved to be slightly higher, viz. 11·9 per cent., this being intermediate between 11·6 per cent. for grain and 13·8 per cent. for straw.

We may next proceed to calculate the standard deviation and probable error for plots of various sizes, using, however, this time an abbreviated method for obtaining the standard deviation which is described in a footnote[1]. We thus obtain a standard deviation of

---

[1] Group the plots as in construction of a frequency curve. Select the group in which the mean is expected to lie (any other group may be taken without altering the result but the arithmetic is generally less arduous if the origin is taken near the mean) and from this as origin index the groups positively and negatively. Multiply the number in each group by the index number of the group. Sum and divide by the total number of observations. This gives ν_1. Next multiply the number in each group by the square of the

11·6 per cent. for the smallest plot of one five-hundredth of an acre, 10 per cent. for one two-hundred-and-fiftieth of an acre, 8·9 per cent. for one-hundred-and-twenty-fifth, 6·3 per cent. for one-fiftieth of an acre, 5·7 per cent. for one-twenty-fifth, and 5·1 per cent. for one-tenth of an acre. These figures are gathered together in Fig. 9 and are compared with a theoretical curve calculated from the standard deviation of one five-hundredth acre, on the assumption that the material is homogeneous and subject to no systematic error. It will be seen that the standard deviations are larger for wheat than for mangolds when plots of the same size are compared, a difference which may be due to the crop and the nature of the soil in the two cases.

The graph also shows a very sharp change of curvature at about one-fiftieth of an acre, from which we may conclude that there is little

index number. Sum, and divide by the total number of observations. This gives ν_2. Then S. D. $= \sqrt{\nu_2 - \nu_1^2 - \frac{1}{12}} \times$ unit of grouping.

In order to compare standard deviations derived from samples of varying size this value may be multiplied by $\sqrt{\dfrac{n}{n-1}}$, but as soon as this begins to make any appreciable difference it is a sign that the sample is growing too small to give an accurate value of the standard deviation.

$\nu_1 \times$ unit of grouping represents the difference between the true mean and the mid-point of the selected group. In case there are only a few observations the group unit should be small or there may be an appreciable difference between the mean and the S. D. as determined from the grouped figures and the actual figures. Where the number is large the error introduced in this way is negligible. For example the mean and S. D. found below from groups of 0·10 are 3·945 and 0·458, while from the same observations from groups of 0·20 they are 3·944 and 0·460.

E.g. 500 wheat plots

3·50 – 3·59	26 × −4 =	−104 × −4 =	416
3·60 – 3·69	37 × −3 =	−111 × −3 =	333
3·70 – 3·79	44 × −2 =	− 88 × −2 =	176
3·80 – 3·89	34 × −1 =	− 34 × −1 =	34
3·90 – 3·99	46 × 0		
4·00 – 4·09	42 × 1 =	42 × 1 =	42
4·10 – 4·19	28 × 2 =	56 × 2 =	112
4·20 – 4·29	41 × 3 =	123 × 3 =	369
4·30 – 4·39	31 × 4 =	124 × 4 =	496
	Total	−2	10542

$$\text{S. D.} = \sqrt{\frac{10542}{500} - \left(\frac{-2}{500}\right)^2 - \frac{1}{12}} \times \cdot 1 \times \sqrt{\frac{500}{499}} = \cdot 458.$$

$$\text{Mean} = 3·945 - \tfrac{2}{500} \times \cdot 1 = 3·945.$$

We have not grouped the observations in any case where the number of plots was less than 100.

advantage to be obtained by increasing the size of plots of wheat above one-fiftieth of an acre.   We may again try the effect of scattering the plots.   Making up plots of one-hundredth of an acre from five scattered plots each one five-hundredth of an acre, we obtain a standard deviation of 4·6 per cent.   Doubling the size of the plots and making up one-fiftieth of an acre in five scattered plots, each one-two-hundred-and-fiftieth of an acre, the standard deviation becomes 3·6 per cent. This is reduced to 2·4 per cent. when the area is made up to one-tenth of an acre in five scattered fiftieths, and it is only reduced to 1·97 per cent. when the plot is made one-fifth of an acre in five scattered twenty-fifths.

To test the number of scattered units into which it is most economical to divide a given area, the standard deviation of one-fifth acre plots made up in various ways was calculated.   The results are set

TABLE IX.   *Standard deviation in one-fifth acre plots composed of different numbers of units; wheat.*

No. of units	Standard Deviation per cent.
1	4·60
2	2·55
5	1·97
10	1·57

out in Table IX, and confirm the fact that it is undesirable to increase the number of duplicates beyond four or five.

Thus we confirm the evidence afforded by the standard deviation graph that the probable error is reduced sufficiently near to the minimum when the plots are one-fiftieth of an acre in area, beyond which point increase of size does not secure much greater accuracy. From five scattered plots each one-fiftieth of an acre in area we may expect to get a result with a probable error of about 1·6 per cent., or with a 40 bushel yield the odds are even that the weight from such an experimental plot will lie within 39·4 and 40·6 bushels.

It will be seen that these results for wheat agree with those previously obtained for mangolds, that the error attached to a single plot cannot be greatly reduced by increasing its area above one-fortieth or one-fiftieth of an acre, and that the probable error can be reduced to a working minimum by taking four or five similar plots scattered about

the field under experiment. We are indebted to "Student," by whose assistance and criticism we have been greatly aided in the whole of this discussion of our experimental results, for the working out of a method whereby the experimental error may be still further reduced when only a single comparison is desired, as for example between two varieties or two methods of manuring, by taking advantage of the correlation which exists between adjacent areas. This contribution is set out in an Appendix. Meantime the following general conclusions may be drawn from our results:

(1) In all field trials, however uniform the land and careful the management of the experiment, there will be an experimental error attached to the result, due to so many incalculable factors that it may be described as casual.

(2) The error diminishes with the size of the plot, but the reduction is small when the plot grows above one-fortieth of an acre.

(3) The error may be best diminished by increasing the number of plots similarly treated and scattering them about the area under experiment, but there is not much to be gained by increasing the number of plots above five.

(4) For practical purposes the authors recommend that in any field experiment each unit of comparison (variety, method of manuring, etc., according to the subject of the experiment) should be given five plots of one-fortieth of an acre each, systematically distributed within the experimental area.

This will reduce the experimental error to within two per cent. of the result, if the land is at all suited for experiment; it does not however eliminate variations due to the unequal effects of different seasons upon the varieties or the action of the manures under experiment. Such variations can only be eliminated by continuing the experiments for several years. Similarly variations induced by the type of soil can only be ascertained by repeating the experiments on several soils.

APPENDIX.

NOTE ON A METHOD OF ARRANGING PLOTS SO AS TO UTILISE A GIVEN AREA OF LAND TO THE BEST ADVANTAGE IN TESTING TWO VARIETIES.

THE authors have shown that to reduce the error as low as possible it is necessary to "scatter" the plots. I propose to deal with this point in the special case when a comparison is to be made between only two kinds of plots, let us say two varieties of the same kind of cereal.

If we consider the causes of variation in the yield of a crop it seems that broadly speaking they are divisible into two kinds.

The first are random, occurring at haphazard all over the field. Such would be attacks by birds, the incidence of weeds or the presence of lumps of manure. The second occur with more regularity, increasing from point to point or having centres from which they spread outwards; we may take as instances of this kind changes of soil, moist patches over springs or the presence of rabbit holes along a hedge.

Having made this distinction between random and regular causes of variation let me hasten to add that almost all causes of variation may belong to one or other or both of these classes according to the size of the plot in question.

In any case a consideration of what has been said above will show that any "regular" cause of variation will tend to affect the yield of adjacent plots in a similar manner; if the yield of one plot is reduced by rabbits from a bury near by, the plot next it will hardly escape without injury, while one some distance away may be quite untouched and so forth. And the smaller the ·plots the more are causes of variation "regular"; for example, with large plots a thistly patch may easily occur wholly within a single plot leaving adjacent plots nearly or altogether clean, but with quite small plots one which is overgrown with thistles is almost sure to have neighbours also affected.

Now if we are comparing two varieties it is clearly of advantage to arrange the plots in such a way that the yields of both varieties shall be affected as far as possible by the same causes to as nearly as possible an equal extent.

To do this it is necessary, from what has been said above, to

compare together plots which lie side by side and also to make the plots as small as may be practicable and convenient.

There is a reason, apart from the difficulty of cultivating very small plots, why the plots should not be made too small and that is, that when two different varieties are sown next one another the outside drill of each is under abnormal conditions and if it be counted in the plot may introduce an error which in a small plot may be quite substantial, but if it is not counted the space wasted by rejecting the outside drills of small plots becomes considerable.

Let us suppose that the smallest practicable size of plot has been chosen and the land available for the comparison has been divided up into plots of this size and sown, chequer fashion, with seed of the two varieties.

Obviously nothing that we can do (supposing of course careful harvesting) can now alter the accuracy of the resulting comparison of yields, but we can easily make different estimates of the reliance which we can place on the figures.

For example, the simplest way of treating the figures would be to take the yields of the plots of each variety and determine the standard deviation of each kind. Then from published tables we can judge whether such a difference as we find between the total yields is likely to have arisen by chance.

An advance on this is to compare each plot with its neighbour and to determine the standard deviation of the differences between these pairs of adjacent plots.

From what has been said above as to the occurrence of "regular" sources of error it will be seen that such differences as these will be to a much larger extent dependent on the variety, and to a less extent on errors, than if the mere aggregates are compared.

The standard deviation will therefore be smaller and the confidence which can be placed in the result increased.

By a further device we can still further decrease the standard deviation and increase our certainty.

For if, instead of harvesting the whole of each plot together, we divide each plot into two before harvesting (and that this can be done is clear from the account of the work done with the mangolds and wheat), then we get twice the number of comparisons, and the plots being half the size are comparatively closer together and the error of their comparison is reduced.

But, it will be asked, why take all this trouble? The error of

comparing plots of any given size has been found by the authors of the paper, and all that has to be done is to apply this knowledge to the particular set of experiments.

The answer to this is that there is no such thing as the absolute error of a given size of plot.   We may find out the order of it, be sure perhaps that it is not likely to be less than (say) five per cent. nor more than 15 per cent. without producing visible heterogeneity, but the error of a given size of plot must vary with all the external conditions as well as with the particular crops upon which the experiment is being conducted, and it is far better to determine the error from the figures of the experiment itself; only so can proper confidence be placed in the result of the experiment.

The diagram illustrates the proposed method of arranging the plots.
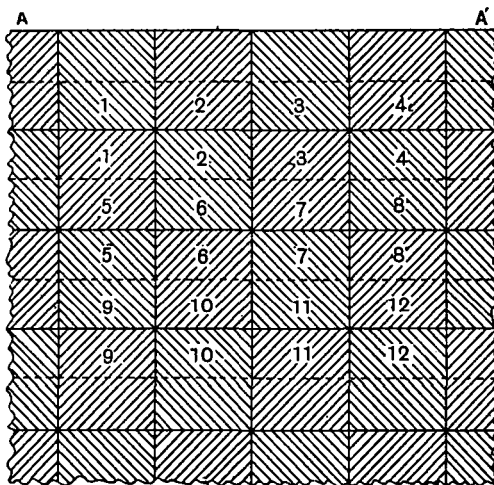
The different shading represents the two different varieties.

The firm lines represent the outside of the original plots.

A A' is part of the boundary of the experimental ground, part of which is given in the diagram.

The dotted lines show the further division made at harvesting.

Then the yields of the half-plots 1, 1 : 2, 2 : ... etc., are compared together.

The outside half-plots are neglected as it is usual to discard the edge of the field.

I have determined the error of comparing plots of different sizes in this way both with the mangold and the wheat figures.

Considering first the mangolds:

The crop on half an acre in the present experiment was about 32,860 lbs., and the standard deviation of a single one-two-hundredth acre was found to be 20·37 lbs. Hence the standard deviation of half an acre made up at random from 100 such small plots would be 20·37 × √100 or 203·7 lbs., and the standard deviation of the comparison between two such half-acres would be 203·7 × √2 or 287 lbs.

This would amount to ·87 per cent. so that one could not begin to be sure that a difference between two varieties of mangolds compared in this way (one-two-hundredth plots arranged at random) until it amounted to say 2·6 per cent.

But now suppose that the plots were each originally one-hundredth acre, bisected at harvest and compared as suggested above.

Then the actual figures given by the authors enable us to determine the standard deviation of the difference between the half-acre.

It amounts to no more than 223 lbs. or ·68 per cent. *I.e.* although working with plots twice the size up to harvest time we get the same accuracy with one acre of ground as would have been obtained with $\left(\dfrac{·86}{·68}\right)^2$ acres or 1·65 acres on the first plan.

Now suppose the plots to be one-fiftieth divided into one-hundredths at harvest.

Then I find the S. D. to be 274 lbs. or ·83 per cent.

Similarly $\frac{1}{25}$th acre plots harvested as $\frac{1}{50}$ths give a S. D. of comparison 289 lbs. or 0·88 %

,, $\frac{1}{10}$,, ,, square $\frac{1}{50}$,, ,, 374 ,, 1·14 %

,, $\frac{1}{10}$,, ,, long $\frac{1}{50}$,, ,, 329 ,, 1·00 %

With such small numbers the difference between the last two cannot be taken as significant, but one would expect the square plot to give a worse comparison than the long plot.

We may summarise the above results in the table below.

Size of plot	Percentage S. D. of comparing $\frac{1}{2}$ acres	Total area required to give a S. D. of 1 % in the comparison
$\frac{1}{100}$th harvested as $\frac{1}{200}$ths	·68	·46 acre
$\frac{1}{50}$,, ,, $\frac{1}{100}$,,	·83	·69 ,,
$\frac{1}{25}$,, ,, $\frac{1}{50}$,,	·88	·77* ,,
$\frac{1}{10}$,, square $\frac{1}{50}$,,	1·14	1·30* ,,
$\frac{1}{10}$,, long $\frac{1}{50}$,,	1·00	1·00* ,,

* These samples are too small to give more than a rough indication of the S. D. and of the area required. I have elsewhere (*Biometrika*, VI, p. 19) given special tables for dealing with such small numbers.

The corresponding figures derived from the wheat results are set out in the second table:

Size of plot	S. D. in lbs. of comparing two half-acres	S. D. as a % of crop on a half-acre	Total area required to give a S. D. of 1 % in the comparison
$\frac{1}{100}$th divided into $\frac{1}{500}$ths at harvest	7·02	·71	·50 acre
$\frac{1}{125}$,, $\frac{1}{750}$,, ,,	8·54	·86	·74 ,,
$\frac{2}{125}$,, $\frac{1}{125}$,, ,,	11·56	1·17	1·37 ,,
$\frac{1}{25}$,, $\frac{1}{50}$,, ,,	10·40	1·05	1·10* ,,
$\frac{1}{25}$,, $\frac{1}{25}$,, ,,	19·40	1·96	3·84* ,,
$\frac{1}{100}$ taken at random	10·28	1·04	1·08 ,,

* These samples are too small to give more than a rough indication of the S.D. and of the area required. I have elsewhere (*Biometrika*, vi, p. 19) given special tables for dealing with such small numbers.

Both these tables show that in the actual fields which were measured, the area of land required to give a comparison between two varieties would increase rapidly as the size of plot increased if the same accuracy were required in the result.

Roughly speaking one-twentieth acre plots of mangolds would require at least twice as much land as one-two-hundredth acre plots in order that we may place as much confidence in the result, while one-fiftieth acre plots of wheat would probably require more than twice as much as one-five-hundredth acre plots.

Hence it is clearly of advantage to use the smallest practicable size of plot.

Also the advantage of comparing adjacent plots is apparent in these examples, since with the roots less than two-thirds of the land is required to give the same accuracy as random comparison and with the wheat less than half.

Of course the comparison of whole half-acre plots would be liable to give errors of quite a different order: thus the South half-acre of mangolds is 4·7 per cent. better than the North half-acre, while the West half-acre of wheat is 8·3 per cent. better than the East half-acre; such differences would be quite impossible if the half-acres were sub-divided into the smaller sizes of plots.

<div align="right">STUDENT.</div>