Sanders, H. G. 1930. A note on the value of uniformity trials for subsequent experiments. *The Journal of Agricultural Science.* 20 (1), pp. 63-73.

# A NOTE ON THE VALUE OF UNIFORMITY TRIALS FOR SUBSEQUENT EXPERIMENTS.

By H. G. SANDERS, M.A., Ph.D.

(*From the Statistical Department, Rothamsted Experimental Station, Harpenden.*)

It is now generally accepted that the soil of any field may be assumed to be markedly variable, and that, in consequence, treatments or varieties must be replicated in experimental trials, so that the error caused by soil differences may be estimated, and the significance of the results appreciated: variations in fertility over the area covered are always found in uniformity trials—that is to say, trials in which the whole area was treated exactly similarly, but plots in which were harvested, and the produce weighed, separately. In view of these irregularities, one possibility of increased precision appears to lie in carrying out a preliminary uniformity trial, harvesting the crop separately in the various plots to be used for a subsequent experiment, and so mapping out the fertility of the field by the plot yields when all are treated alike: these figures could be used later to correct the yields of the plots when under experiment, and so to circumvent to a certain extent the field experimentalist's bugbear—soil heterogeneity: it is conceivable that in this way a degree of precision might be attained with the actual experiment which could only be reached otherwise by a greater number of replicates than could well be managed, and that consequently the extra labour involved in preliminary uniformity trials might be justified. Naturally the possibilities depend entirely on the constancy of the plots in their relative productivity in different years, under different climatic conditions, and, usually, under other crops: this question could readily be explored were figures available giving the results of uniformity trials carried out over a series of years, on the same fields, sub-divided into the same plots, but such data are rare, and those discussed here are recognised as inadequate to give anything approaching a final solution. It was thought, however, that a note might serve to direct attention to the problem, and might stimulate others, who have access to suitable and more abundant data, to investigate them with the same end in view: it was also felt that some workers might be glad of an example of how the results of a uniformity trial might be used in subsequent experiments on the same plots, so that the method of working is described in some detail, at the risk of appearing obvious.

THE METHOD OF USING THE INFORMATION.

If the yield of any plot under a preliminary uniformity trial (or its mean yield under a plurality of trials) be denoted by $x$, and its yield in the actual subsequent experiment by $y$, then some method of correcting $y$ for $x$ must be selected. At first sight the simplest procedure would appear to be to take the difference between the two—the efficacy of any one experimental treatment being tested by the mean value of $y - x$ for the plots on which it was carried out: if $x$ and $y$ were expressed as percentages of the mean yield of all plots in their respective years, then a positive value of $y - x$ would indicate that that particular treatment had been beneficial relative to the other treatments employed, and vice versa. Statistically, however, this method may lead to a loss, rather than a gain, in precision, for if $V_x$ be the variance of $x$, then

$$V_{(y-x)} = V_y + V_x - 2r_{xy}\sqrt{V_x V_y}.$$

Consequently if there is no correlation between the plot yields in the two years, this method would add to the variance of $y$ that of $x$: with $V_x$ and $V_y$ approximately equal, no gain at all would be made unless $r_{xy}$ exceeded $+0{\cdot}5$, whilst if it were negative there would be a serious loss.

A gain may, however, be effected by means of the regression function between the two variables: if $V_{y \cdot x}$ denote the variance of $y$ corrected for $x$ by means of this (or, as it is sometimes stated "with $x$ held constant"), then

$$V_{y \cdot x} = V_y(1 - r^2_{xy}) = V_y - \frac{(\mathrm{Cov}_{xy})^2}{V_x},$$

where $\mathrm{Cov}_{xy}$ is the covariance between $x$ and $y$, or the mean product of their deviations from their means. In this case there is a definite gain if $x$ and $y$ are at all closely related (positively or negatively)—that is, if the produce of an individual plot in one year is any guide to its performance in another. It will be realised that $V_{y \cdot x}$ gives the variance of $y$ corrected by the regression equation $y = bx$, where $b$ is equated to $\mathrm{Cov}_{xy}/V_x$: accordingly it must not be used to test the significance of the difference between the actual mean plot yields given by any two treatments, but that between the mean values of $y - bx$ given by them: if $n$ be the number of replicates, the figures compared will be $S\left(\frac{y - bx}{n}\right)$, or

$\frac{1}{n}(Sy - bSx)$: since, with a random distribution of the treatments over the plots, the values of $Sx$ will tend to be constant, it follows that the differences between these figures will be of the same order of magnitude as the differences between the means of $y$, and consequently any reduction effected in calculating $V_{y \cdot x}$ from $V_y$ will be a direct gain in precision. It can readily be seen that in all circumstances the variance of $y - bx$ will be less than that of $y - x$, for

$$V_{y \cdot x} < V_{(y-x)},$$

if $$V_y (1 - r^2_{xy}) < V_y + V_x - 2r \sqrt{V_x V_y},$$

if $$(\sqrt{V_x} - r \sqrt{V_y})^2 > 0,$$

which is necessarily true, since a square cannot be negative.

## THE DATA.

Uniformity trials were carried out between 1906 and 1911 on two fields at Aarslev (Denmark)(1), which provide (very limited) data to test the possibilities. One field ($A_2$) was divided into 30 plots—6 strips of 5—which were all treated alike but harvested separately, and the crops grown were: 1907, oats; 1908, rye; 1909, barley; 1910, mangolds; 1911, barley. The other field ($E_2$) was divided into 128 plots—16 strips of 8—and carried oats in 1906, barley in 1907, "seeds" in 1908, and rye in 1909: there was a remarkable oscillation in fertility across this field in one direction, the 1st, 3rd, ... 15th strips consistently giving much higher yields than the 2nd, 4th, ... 16th strips—in fact in the four years the odd numbered strips gave a total yield of 27,817, as compared to 23,383 for the even numbered strips. This oscillation apparently arose as a legacy of the old practice of ploughing in high ridges: the tops of the ridges exhibited greater fertility than the borders of the furrows, so that soil was worked from the former to the latter and the field levelled out: this meant that over the site of the old furrows there was a good depth of rich soil, whilst it was very shallow where the ridges had been. The strips were so arranged as to cover the site of the furrow and of the ridge alternately, with the result noted above: in order to escape this variation, the table was condensed by taking 2 strips together (so that the new strips each included the whole of one of the old "lands") making it an 8 by 8 square. The paper referred to above gives the production for each individual plot as a percentage of the mean yield over the field in that year: the method of enquiry was to suppose that the last year for which figures are given was the actual experiment, and to calculate how far the knowledge gained by one, two, three, or four, preliminary uniformity trials would be effective in reducing the residual variance (that is, what remained after the variance that could legitimately be taken out of the total by suitable methods of "local control"), by which the significance of the results would be judged.

There appeared to be two printer's errors in the paper from which the figures were taken: with Field $A_2$ the yields given for the year 1908 add up to 3010 instead of 3000: reference to the Fig. 6 given there seemed to indicate that the excess lay in row 3 and eventually it was decided to reduce plots 3 *c* and 3 *f* to 96 and 84 respectively: again, with Field $E_2$ in 1908, column 10 sums to 791 instead of 786 as shown: reference to Fig. 13 indicated that the yield of plot 10 *g* should probably have been 92 instead of 97. These two slight changes in the data could, of course, have no appreciable effect on the results.

*Field* $A_2$. The following table shows the arrangement of the plots on this field, together with their yields in the supposedly experimental year (1911): the thick lines divide the area into 5 blocks of 6 plots each, suitable for testing six different treatments. These would be assigned to the particular plots wholly at random, with the one restriction that each treatment must occur once in each block: the letters show an arrangement arrived at in this way with treatments denoted by the letters A to F—obtained by entering the letters separately on to six cards, and then shuffling them thoroughly for each block, and writing them down in the order in which the cards were found: since these are dummies they are left out in the working, and only referred to later as an example of the way in which previous information would be used.

| | | | | | |
|---|---|---|---|---|---|
| D<br>103 | A<br>101 | B<br>95 | D<br>104 | C<br>109 | F<br>115 |
| B<br>98 | C<br>102 | E<br>97 | A<br>111 | D<br>109 | E<br>104 |
| F<br>94 | E<br>106 | F<br>96 | C<br>95 | B<br>112 | A<br>94 |
| C<br>90 | B<br>111 | E<br>100 | C<br>96 | F<br>99 | B<br>86 |
| D<br>95 | A<br>105 | F<br>101 | A<br>90 | E<br>98 | D<br>84 |

Using the methods evolved by Dr Fisher (2), the variance in this year was analysed in the following way:

Sum of the squares of the yields on 30 plots = 301,730.
Sum of the squares of the 5 block totals = 1,804,082.

From these the sums of the squares of the deviations were calculated as

$$\text{Total} = S\,(y - \bar{y})^2 = 301{,}730 - 30 \times 100^2 = 1730.$$

$$\text{Between blocks} = 6S\,(\bar{y}^1 - \bar{y})^2 = 6 \times \frac{1}{6^2}\,(1{,}804{,}082 - 5 \times 600^2) = 680{\cdot}333,$$

giving the following analysis:

| | Degrees of freedom | Sum of the squares | Variance | S.D. | $\text{Log}_e$ S.D. |
|---|---|---|---|---|---|
| Between blocks | 4 | 680·333 | 170·083 | 13·0416 | 2·5683 |
| Within blocks | 25 | 1049·667 | 41·987 | 6·4798 | 1·8687 |
| Total | 29 | 1730·000 | 59·655 | 7·7237 | — |

If, then, 1911 stood by itself the variance of 1 plot due to experimental error would be 41·987, and the significance of the differences between means of 5 replicates would be judged by comparing them to $\sqrt{\dfrac{2 \times 41 \cdot 987}{5}}$. It is seen that the restriction imposed by the block method of local control has been successful in reducing the variance from 59·655 to 41·987: comparing the variance between blocks with that within blocks, $z = 0{\cdot}6996$, which, with $n_1 = 4$, $n_2 = 25$, gives a value of $P$ lying between 0·01 and 0·05, showing that there were real differences between the average fertility of the blocks.

Treating the other years similarly, the variances within the blocks were found to be as follows:

| | |
|---|---|
| 1910 ... ... ... ... ... | 82·947 |
| Mean, 1909–1910 ... ... ... | 29·343 |
| „ 1908–9–10 ... ... ... | 37·785 |
| „ 1907–8–9–10... ... ... | 28·705 |

The sum of the products of the yields in any of these and the yields in 1911 can be found exactly similarly to the sum of the squares in any one year, and, having subtracted that between blocks from the total, and divided by 25, the following covariances were found within the blocks:

| | |
|---|---|
| 1910 × 1911 ... ... ... | + 38·020 |
| 1909–10 × 1911 ... ... ... | + 25·133 |
| 1908–9–10 × 1911 ... ... ... | + 31·053 |
| 1907–8–9–10 × 1911 ... ... | + 26·347 |

Substituting the 1910 ($x$) and 1911 ($y$) figures in the formula given above, we obtain

$$V_{y\cdot x} = V_y - \frac{(\mathrm{Cov}_{xy})^2}{V_x} = 41{\cdot}987 - \frac{(38{\cdot}020)^2}{82{\cdot}947} = 24{\cdot}560.$$

If, however, yields are to be corrected by a linear regression, one more degree of freedom is used up, and consequently we have, finally

$$V_{y\cdot x} = 24{\cdot}560 \times \frac{25}{24} = 25{\cdot}583.$$

Thus the variance of $y - bx$ is some 40 per cent. less than that of $y$: the precision obtained by correcting for the 1910 yields can be compared to that of 1911 alone by considering the number of replicates that would be necessary to reduce the variance of the mean to the same point in each case: these are clearly in the proportion 25·583 : 41·987, that is as 1 : 1·642.

The full results obtained with this field were as follows:

| Trial years ($x$) | $V_{y\cdot x}$ | Relative precision | Gain in precision |
|---|---|---|---|
| — | 41·987 | 1·000 | — |
| 1910 | 25·583 | 1·642 | 0·642 |
| 1909–10 | 21·311 | 1·969 | 0·327 |
| 1908–9–10 | 17·152 | 2·445 | 0·476 |
| 1907–8–9–10 | 18·547 | 2·262 | −0·183 |

There is a progressive lowering of the value of $V_{y\cdot x}$ as one, two and three previous years are taken into consideration, until a point is reached at which the experiment is nearly $2\frac{1}{2}$ times as exact as if no previous uniformity trials had been carried out: the inclusion of 1907 does not, however, improve on this, giving, in fact, a slightly higher value. The year 1909 is of some special interest as then the field carried the same crop (barley) as in the supposedly experimental year: taking only that one into consideration $V_{y\cdot x}$ was found to be 35·904, so that a different crop in the preceding year was a much better basis for correction than the same crop two years before.

*Field* $E_2$. By putting 2 strips together as mentioned above, the plots in this field were reduced to 64, and were in the form of an 8 by 8 square: with such plots an experiment might be set out in the form of a Latin square, or simply by dividing into 8 blocks of 8 plots each, and the efficacy of previous uniformity trials was tested under both of these arrangements. Taking the Latin square first, the simple addition of the squares of the 1909 yields gave the following:

| | |
|---|---|
| 64 plots ... ... ... ... | 2,566,204 |
| Totals of 8 columns ... ... | 20,517,636 |
| Totals of 8 rows ... ... ... | 20,482,768 |

Then, since the mean plot yield was 200 and the mean total yield of rows and columns 1600, we have the following as the contributions to the sum of the squares:

$$\text{Total} = 2{,}566{,}204 - 64 \times 200^2 = 6204$$

$$\text{Columns} = \tfrac{1}{8}(20{,}517{,}636 - 8 \times 1600^2) = 4704{\cdot}5$$

$$\text{Rows} = \tfrac{1}{8}(20{,}482{,}768 - 8 \times 1600^2) = 346$$

giving the following analysis

| | Degrees of freedom | Sum of the squares | Variance | S.D. | $\text{Log}_e$ S.D. |
|---|---|---|---|---|---|
| Between columns | 7 | 4704·5 | 672·071 | 25·9244 | 3·2553 |
| Between rows | 7 | 346·0 | 49·429 | 7·0306 | 1·9503 |
| Error | 49 | 1153·5 | 23·541 | 4·8519 | 1·5794 |
| Total | 63 | 6204·0 | 98·476 | 9·9235 | — |

The effect of correcting $y$ for $x$ by linear regression on the variance ascribable to experimental error, in the case of this arrangement on this field, was as follows:

| Trial years ($x$) | $V_{y \cdot x}$ | Relative precision | Gain in precision |
|---|---|---|---|
| — | 23·541 | 1·000 | — |
| 1908 | 23·129 | 1·018 | 0·018 |
| 1907–8 | 23·737 | 0·992 | − 0·026 |
| 1906–7–8 | 23·981 | 0·981 | − 0·011 |

It is clear that in this field the plots did not tend to keep their relative yield positions from year to year, so that uniformity trials could serve no useful purpose to correct the yields under experiment: there was in fact a slight loss in precision in the last two cases, which arose by the elimination of one more degree of freedom, for if the covariance is zero, $V_{y \cdot x}$ will obviously be $V_y \times \frac{49}{48}$.

In the above case a large part of the sum of squares was taken out in the columns, so that it would appear possible that previous information might have been valuable if the experiment had been planned on the block system, though that design itself would not have been so efficient: with 8 blocks of 8 plots each, however, there was again no reduction of variance by correction for previous yields:

| Trial years ($x$) | $V_{y \cdot x}$ | Relative precision | Gain in precision |
|---|---|---|---|
| — | 42·571 | 1·000 | — |
| 1908 | 41·106 | 1·035 | 0·035 |
| 1907–8 | 42·088 | 1·011 | − 0·024 |
| 1906–7–8 | 42·989 | 0·990 | − 0·021 |

It will be noticed that the variances are much larger in this case, showing that the Latin square method of local control would be much more effective than that of blocks—though 2 strips (2 "columns") were taken together there yet remained a large variation across that direction.

With the original 128 plots of this field it is unlikely that any experiment would have been laid down with no local control, either by strips or by blocks: if, however, treatments had been completely randomised over the whole field, so that the fertility oscillation in the columns could not have been taken out, the experiment would have lost much in precision, but a considerable proportion would have been retrieved by taking the regression on previous uniformity trials into account: for all the deviations caused by soil variations would have gone into error, and, as the relative yields of the plots in the odd numbered columns were consistently high, whilst those in the even numbered columns were low, a large part of that variance would be eliminated in calculating $V_{y \cdot x}$ from $V_y$. This serves to emphasise the point that uniformity trials will be more effective in increasing precision in subsequent experiments on very irregular soils, and where local fluctuations are not adequately controlled by the design of the experiment.

### Experimental trial.

Although preliminary uniformity trials would be futile on Field $E_2$ (that is, if reasonable experimental methods were adopted), they would give valuable information on Field $A_2$, as it appears that there were variations between individual plots (as distinct from those between blocks or strips) of a more or less constant nature—though as to whether this was a matter of plant food, soil texture, drainage, etc., we have no information. The greatest reduction of the variance due to error was given by correction from the yields in the previous three years, so that it was thought desirable to see exactly how the inclusion of this regression would affect the analysis: the random distribution of six dummy treatments is shown on p. 66. The regression of $y$ on $x$ is given by $\frac{\mathrm{Cov}_{xy}}{V_x}$, which in this case takes the value $\frac{38{\cdot}682}{41{\cdot}84645}$ or 0·92438. The various treatments gave:

| | Mean plot yield ($y$) | Mean yield of same plots in 1908–9–10 ($x$) | $(y-bx)$ |
|---|---|---|---|
| A | 100·2 | 98·2 | 9·4 |
| B | 100·4 | 98·8 | 9·1 |
| C | 98·4 | 100·47 | 5·5 |
| D | 99·0 | 100·2 | 6·4 |
| E | 101·0 | 103·8 | 5·0 |
| F | 101·0 | 98·53 | 9·9 |

For actual yields the greatest difference is between E or F and C, and amounts to 2·6: with corrected yields it is seen that E was favoured by the soil allotted to it, and that it actually evoked the least response, differing from F by 4·9.

Neglecting, at first, the regression, the analysis of variance (of $y$) is as follows:

| | Degrees of freedom | Sum of the squares | Variance | S.D. | $\text{Log}_e$ S.D. |
|---|---|---|---|---|---|
| Between blocks | 4 | 680·333 | 170·083 | 13·0416 | 2·5683 |
| Between treatments | 5 | 28·800 | 5·760 | 2·4000 | 0·8755 |
| Error | 20 | 1020·867 | .51·0435 | 7·1445 | 1·9663 |
| Total | 29 | 1730·000 | 59·655 | 7·7237 | — |

The random distribution of treatments has been peculiar in giving such small differences: comparing the treatment variance with that due to error, $z = -1·0908$, which, with $n_1 = 20$ and $n_2 = 5$, gives a value of $P$ lying between 0·01 and 0·05, showing that such evenness would be given by chance less than 5 times in 100: since practically nothing is taken out by treatment, and 5 degrees of freedom are sacrificed to it, the variance due to error is raised.

The variance of $(y - bx)$ may be determined directly by preparing a fresh table, giving instead of the experimental yield of each plot the value of $(y - bx)$, and by obtaining and analysing the sum of the squares from these new figures. Alternatively it may be derived from tables already calculated for the analysis of the variance of $y$, and of $x$, and of the covariance of $xy$: each of these will contain four rows—one showing the variance (or covariance) between blocks, one between treatments (these of course will always be purely hypothetical in the case of $x$), one for experimental error, and one for the total. For the compound observation $(y - bx)$ we have

$$S\,(y - bx)^2 = S\,(y^2) - 2bS\,(xy) + b^2S\,(x^2),$$

so that the new table can be constructed by applying this formula to the other three tables, row for row all through. Since $(y - bx)$ contains one statistic ($b$) already obtained from the data, the total number of degrees of freedom will be reduced from 29 to 28, and, as $b$ has been calculated from the figures for error, this one degree of freedom will be taken from that row. In this way the following analysis of the variance of $(y - bx)$ was obtained:

| | Degrees of freedom | Sum of the squares | Variance | S.D. | $\text{Log}_e$ S.D. |
|---|---|---|---|---|---|
| Between blocks | 4 | 531·133 | 132·783 | 11·5230 | 2·4444 |
| Between treatments | 5 | 115·854 | 23·171 | 4·8136 | 1·5715 |
| Error | 19 | 305·727 | 16·0909 | 4·0113 | 1·3890 |
| Total | 28 | 952·714 | 34·0255 | 5·8331 | — |

In this case our random distribution of treatments shows no peculiarities, the treatment variance agreeing closely with that due to error.

The comparative efficiency in the two cases can be seen from the following:

| | Without regression | With regression |
|---|---|---|
| Variance of 1 plot ... ... | 51·0435 | 16·0909 |
| Variance diff. of means of 5 plots | 20·4174 | 6·4364 |
| S.E. ... ... ... ... | 4·5186 | 2·5370 |

A table of $t$ shows that $P = 0{\cdot}01$ when $t = 2{\cdot}845$ with 20 degrees of freedom, and when $t = 2{\cdot}861$ with 19 degrees of freedom, so that a difference of 12·86 would be needed for this standard of significance where the regression was not introduced, but only one of 7·26 when that part of the variance accounted for by it was eliminated. As shown by the values of $z$ in the two cases, the former is not approached by any of the differences between the treatment means, neither is the latter by any of the differences between the mean values of $(y - bx)$: this was to be expected since the treatments were dummies, but had real treatments been carried out it will be seen that the use of the regression would have materially increased the exactitude of the comparison—in fact under the latter conditions 6 replicates would have provided as precise information as 19 would have done if the experimental year gave our only measure of the productivity of the plots. This present insertion of hypothetical treatments into the data shows a greater increase in precision than indicated by the value of $V_{y \cdot x}$ on p. 68, where the relative accuracy in the case of these three preliminary trials was shown to be as $1 : 2{\cdot}445$: the discrepancy arises in the chance distribution of the treatments over the plots, and the value given there for $V_{y \cdot x}$ is the more definite, for that shows what would be the average figure, if a large number of trials with dummy treatments, such as the present, were carried out.

In this instance the gain is very considerable, and it is possible that under certain circumstances (*e.g.* with a restricted area, or where little assistance was available at any one time) it might repay the labour of three years uniformity trials, even though it would increase the work fourfold and the precision but little more than three-fold: such a result must not however be expected in all cases, for with Field $E_2$, as shown above, the yields under previous trials would be quite ineffective as a basis for correction. The final decision of this question of the value of uniformity trials in this direction must await the analyses of a number of series of uniformity trials carried out on other fields, and it will vary in each particular case according to the degree of soil heterogeneity of the experimental area, and the adequacy of the methods of local control that are adopted: it is thought however that the present results, meagre though they are, suffice to indicate that in some cases appreciable advances might possibly be made along this line.

## Summary.

The question attacked is whether soil variations are sufficiently constant from year to year to give useful corrections to the yields of experimental plots from their yields under previous uniformity trials, and the data investigated were the published results of uniformity trials carried out on two fields at Aarslev (Denmark) between 1906 and 1911. In one case the plots did tend to keep constant in their relative yields, and the precision of an experiment would be increased by nearly 150 per cent. if the regression on the mean yield in the three previous years were used: with the other field, however, the plots showed no constancy in yield (when the variation due to strips was taken out as in modern experimental methods), and consequently previous uniformity trials could give no assistance.

The work described here was done whilst I was enjoying the hospitality of the Rothamsted Laboratory: it gives me much pleasure to acknowledge my great indebtedness to Dr R. A. Fisher, F.R.S., who suggested the problem, and guided my unsteady steps throughout.

## REFERENCES.

(1) Hanson, N. A. "Prøvedyrkning paa Forsøgsstationen ved Aarslev." *Tids. for Landbrugets Planteavl* (1914), **21**, 553.

(2) Fisher, R. A. *Statistical Methods for Research Workers* (1928). Edinburgh, 2nd edition.

(*Received September 24th*, 1929.)