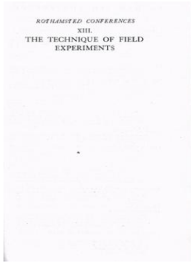
XIII. The Technique of Field Experiments


ROTHAMSTED CONFERENCES
XIII.
THE TECHNIQUE OF FIELD EXPERIMENTS

Full Table of Content

Default Title

Rothamsted Research

Rothamsted Research (1932) *Default Title ;* Xiii. The Technique Of Field Experiments, pp 2 - 64 **- DOI: https://doi.org/10.23637/ERADOC-1-214**

ROTHAMSTED CONFERENCES

XIII.

THE TECHNIQUE OF FIELD EXPERIMENTS

https://doi.org/10.23637/ERADOC-1-214

CONTENTS

PRINTED IN GREAT BRITAIN BY THE ABERDEEN UNIVERSITY PRESS, ABERDEEN

ROTHAMSTED CONFERENCES

The booklets in this series contain the papers and discussions at the conferences held from time to time at Rothamsted on present-day problems in crop production. The papers are written by well-known experts and discussed by some of the best practical farmers. Published by Messrs. Ernest Benn Ltd.

Obtainable from the Secretary, Rothamsted Experimental Station, Harpenden, Herts.

THE GROWING OF LUCERNE

With contributions by Sir JOHN RUSSELL, F.R.S.; H. G. THORNTON, B.A.; A. CUNNINGHAM, B.Sc.; J. MACKINTOSH, N.D.A., N.D.D.; R. D. WILLIAMS, M.Sc. *Price 1s. 6d. net*

THE CULTURE AND MANURING OF FODDER CROPS

With contributions by Sir JOHN RUSSELL, F.R.S.; W. A. C. CARR, M.C.; N.D.A.; J. C. BROWN, P.A.S.I.; C. HEIGHAM, M.A. *Price 1s. 6d. net*

GREEN MANURING

With contributions by H. DREWITT; H. UPCHER; H. INSKIP; H. J. PAGE, M.B.E., B.Sc., A.I.C.; J. A. VOELCKER, M.A., Ph.D.; W. BRUCE, B.Sc.; H. MATTINSON, B.Sc.; A. W. OLDERSHAW, B.Sc. *Price 2s. net*

THE CULTURE AND MANURING OF SUGAR-BEET

With contributions by J. M. VAN BOMMEL VAN VLOTEN; T. G. FOWLER; C. HEIGHAM, M.A.; H. J. PAGE, B.Sc., A.I.C.; I. J. SCHAPRINGER; C. J. CLARKE; R. N. DOWLING, N.D.A., P.A.S.I. *Price 2s. 6d. net*

THE ART AND SCIENCE OF CULTIVATION

With contributions by Sir E. J. RUSSELL, D.Sc., F.R.S.; B. A. KEEN, D.Sc., F.Inst.P.; H. DREWITT; J. STEEL; J. H. SPILMAN; J. JOYCE; C. HEIGHAM, M.A. *Price 2s. net*

POWER FOR CULTIVATION AND HAULAGE ON THE FARM

With contributions by B. A. KEEN, D.Sc., F.Inst.P.; H. C. BURFORD, M.I.A.E., M.I.Mech.E.; G. W. WATSON, M.I.A.E., M.I.Mech.E.; E. PORTER, B.Sc. (Agric.); R. D. MOZER; R. BORLASE MATTHEWS, Wh.Ex., M.I.E.E. *Price 2s. 6d. net*

MALTING BARLEY

With contributions by ROBERT V. REID; H. HUNTER, D.Sc.; JAMES STEWART; JOHN JOYCE; G. H. NEVILE; Sir JOHN RUSSELL, D.Sc., F.R.S.; B. A. KEEN, D.Sc., F.Inst.P. *Price 2s. 6d. net*

RECENT CHANGES IN SYSTEMS OF HUSBANDRY IN ENGLAND

With contributions by Sir JOHN RUSSELL, D.Sc., F.R.S.; C. S. ORWIN, M.A.; F. RAYNS, M.A.; Col. G. H. LONG; JOHN PORTER, B.Sc., N.D.A.; J. W. REID, N.D.A.; W. A. STEWART, M.A.; R. N. DOWLING, N.D.A.; H. DREWITT; J. C. WALLACE, M.A.; W. B. MERCER, B.Sc.; H. V. TAYLOR, O.B.E., B.Sc. *Price 2s. 6d. net*

THE HERTFORDSHIRE AGRICULTURAL SITUATION

With contributions by Sir JOHN RUSSELL, D.Sc., F.R.S.; R. J. THOMPSON, C.B., O.B.E.; J. HUNTER-SMITH, B.Sc.; G. DALLAS; A. W. STREET, C.I.E., M.C.; R. R. ENFIELD; F. J. PREWETT, M.A.; D. CRAWFORD *Price 2s. net*

THE GROWTH OF CHEAPER WINTER FOOD FOR LIVE STOCK

With contributions by Sir JOHN RUSSELL, D.Sc., F.R.S.; J. G. STEWART; Dr. H. E. ANNETT; W. BRUNTON; Captain R. STALLARD; HARALD FABER; W. O. WATT; J. R. KEEBLE; T. C. WARD; W. A. C. CARR, M.C., N.D.A. *Price 2s. 6d. net*

THE MAKING OF NEW GRASSLAND: EXPERIENCES OF PRACTICAL FARMERS

With contributions by the Earl de la WARR; Sir JOHN RUSSELL, D.Sc., F.R.S.; Sir DANIEL HALL, F.R.S.; J. CRUICKSHANK; W. M. FINDLAY, N.D.A.; A. McARTHUR; Major J. KEITH; J. ALSTON; W. S. MANSFIELD, M.A.; C. H. GARDNER; Captain A. R. McDOUGAL; M. JONES. *Price 2s. 6d. net*

THE PLACE AND MANAGEMENT OF SHEEP IN MODERN FARMING

With contributions by J. EGERTON QUESTED; F. A. THOMSON, B.Sc.; J. R. WOOD; A. LEWIS; JOHN JOYCE; Major V. S. BLAND; HENRY EDGAR; H. W. DREWITT; A. C. HILL; H. G. MILLER, B.Sc.; N. V. HEWISON; and others. *Price 1s. 6d. net*

FOREWORD

By Sir E. JOHN RUSSELL

Director of the Rothamsted Experimental Station, Harpenden

The papers presented herewith were read at a Conference at Rothamsted on 7th May, when the technical experts concerned in the carrying out of field experiments assembled to describe the methods they are actually adopting.

The old method, involving the use of large single plots placed side by side, is simple and effective for the purpose of demonstrating known facts so long as the differences to be observed are large and do not require to be expressed numerically. But it breaks down as soon as accurate measurements are needed, because it takes no account of variations in the soil, which are now known to be considerable.

In the classical fields at Rothamsted Lawes and Gilbert had two unmanured plots, one at each end of the series: the yields from these usually differed by about 10 per cent. The differences they were demonstrating were, however, much larger, so that they felt safe in attributing their results to the treatments. In 1846-47 they split the Broadbalk plots lengthwise into two halves, which from that time onwards were harvested separately: this was the first duplication of field plots of which we can find any record. In 1847-48 and occasionally afterwards one half of each plot was treated differently from the other, so that they ceased to be strict duplicates.

Better duplication appears to have been practised by P. Nielsen, the founder of the Danish experimental stations about 1870, in his experiments on grass mixtures for pastures. Some of the Norfolk chamber of agricultural experiments carried out in the later 1880's were systematically repeated thus: A B C D D C B A.

Nothing more was done in this country till 1909. In that year A. D. Hall, and somewhat later T. B. Wood, both urged the need for estimating experimental errors in field work and gave approximate methods for doing this. Somewhat later Dr. E. S. Beaven of Warminster designed his well-known strip method of replication

https://doi.org/10.23637/ERADOC-1-214

specially suitable for variety trials and adopted by the National Institute of Agricultural Botany; it is described by Mr. S. F. Armstrong on page 30.

The subject was taken up seriously at Rothamsted in 1919 and has been much developed since. In 1919 Dr. R. A. Fisher applied to the study of variation an arithmetical analysis known as the analysis of variance, which had the advantage over the ordinary calculus of correlations of avoiding both the calculation of a large number of irrelevant values and also the numerous corrections to which correlations are liable, especially with small samples. He applied the method to the Broadbalk wheat yields and showed its value for measuring the effect of distinct groups of causes. This investigation, however, showed the need for more exact methods for treating the small number of cases, or samples, generally available in agricultural investigations. The first example of an analysis of variance in its modern form was the examination of the results of T. Eden's experiment in 1922 on the response of different potato varieties to manures (Fisher and Mackenzie, *Journ. Agric. Sci.*, 1923). Somewhat later, "Student" gave alternative proofs by himself and by Fisher of formulæ appropriate to cereal variety experiments. Thus rigorous methods of statistical examination were elaborated.

The next step was to develop a correspondingly rigorous field technique, and this was done by Dr. Fisher in co-operation with T. Eden and E. J. Maskell. The chief difficulty was to overcome the effects of the irregularities in the soil which had long been a serious stumbling-block to field experimenters.

Part of the irregularity or heterogeneity could be eliminated by suitable arrangements of the plots, but there was always an unknown remainder of residual errors. It was shown that the statistical analysis previously developed could eliminate the effects of soil irregularities and at the same time afford a valid estimate of the remaining errors, provided that the plots were sufficiently replicated and deliberately randomised.

Dr. Fisher then devised various types of experiments to meet the requirements of the statistical analysis, and tested these on the results of uniformity trials so as to discover which were the most accurate and convenient in actual working. Two types stood out as satisfactory; randomised blocks and the Latin square. The randomised block is the simpler and the more easily adjusted to suit the pecularities of the field and the crop. The experimental area is divided into several strips or blocks, each of which contains one plot of each treatment, the arrangement being deliberately at random and determined not by selection, but by writing the possible arrangements on separate cards, shuffling them, and drawing one out. Since one block is not directly compared with another, the differences in soil fertility between them are eliminated; and since the arrangement within the blocks has been entirely at random, the significance of the results can be estimated.

The Latin square is the more accurate but less widely applicable in fertiliser experiments. The plots are arranged with as many rows and columns as there are treatments. Each treatment appears once, and only once, in each row and each column. A surprisingly large number of arrangements are possible, but the selection is again deliberately at random and, as before, is effected by the shuffling and drawing of cards. From the figures for yield, a standard error is worked out which shows the degree of trustworthiness of the result. A difference in yield equal to the standard error of this difference can be obtained about once in three trials, even when the experimenter is convinced that he has given exactly the same manuring and cultivation to each of the plots, but a difference twice this size would be obtained by chance only once in twenty-two times: it is therefore much more likely to be true. The chances against the difference in yield being due to causes other than the difference in treatment are:—

For difference equal to its standard error .	3 to 1
For difference double its standard error .	22 to 1
For difference three times its standard error .	370 to 1
For difference four times its standard error .	15,780 to 1

For most agricultural purposes a chance of about 30 to 1 is good enough. The "standard errors" for the yield values have to be multiplied by 1·414 (*i.e.* $\sqrt{2}$) in order to give the standard error of the difference between treated and untreated plots—the figure one usually wants. To attain a probability of 30 to 1, a difference must be roughly three times the standard error of the yield.

Dr. Wishart (p. 15) shows how the results are to be worked out. Our experience proves that the methods are quite practicable not only on the fields of the Experimental Station but also on those of ordinary farms: Mr. Garner (p. 49) gives particulars of the methods, and Messrs. Lewis, Manson and Proctor (p. 37) show how to extend them to cover a series of trials. Large numbers of these experiments are now made, the numbers of plots in each ranging from 16 to 144. Usual standard errors per plot on our present methods of good working are:—

USUAL STANDARD ERRORS PER PLOT FOR GOOD WORKING

	Weight per Acre		*Per Cent. of Yield*	
	Randomised Block.	*Latin Square*	*Latin Square*	*Randomised Block*
Potatoes . .	0·7 tons	0·4 tons	4·4	8·4
Sugar Beet: Roots	0·8 ,,	0·6 ,,	5·7	10·2
Tops		0·7 ,,	5·7	10·9
Barley: Grain .	1·5 cwt.	1·3 cwt.	5·6	9·1
Straw .	1·9 ,,	1·9 ,,	6·0	7·2
Oats: Grain .		2 ,,	8	
Straw .		2 ,,	6	

The standard error precisely measures the accuracy of the experiment and it includes errors of working, inequalities due to variable natural agencies, such as weather, birds, insects, diseases, and also soil variations within the individual plots, but not the large variations between plot and plot, which are eliminated by the method of arranging the experiment. It is not, however, an absolute measure, since it depends to some extent on the size and arrangement of the plots. Thus a standard error of 0·4 tons per acre of potatoes in a Latin square experiment is not strictly comparable with a standard error of 0·4 tons in a randomised block experiment having more plots. Nevertheless, it is a useful guide to the experimenter, as showing the standard of performance he is attaining in his work. The standard error is much the same whether the crop is large or small, so that a heavy crop has a lower percentage error than a light one.

There are several plots of each treatment, and the standard error of the final result is much less than the figures of errors per plot; it is usually now at Rothamsted about 2 to 4 per cent. of the mean yield.

In addition detailed observations are made on various growth factors: these are discussed by Mr. Watson on page 54.

These principles have been applied to horticultural experiments by Mr. Hoblyn, who deals with the problem on page 42.

CHAIRMAN'S OPENING REMARKS

SIR A. D. HALL

Chairman

IT gives me great pleasure to be present at to-day's conference, because I can claim a sort of paternal interest in the subject. I believe a communication of mine to the British Association, printed in the Journal of the Board of Agriculture for 1909, in which I discussed the probable error attached to a comparison of a pair of plots as deduced from some of the Rothamsted data, was the first presentation of the problem in this country. That led to the harvesting of an apparently uniform area of mangolds and wheat in units of $^1/_{200}$ and $^1/_{500}$ acre respectively by W. H. Mercer and myself, and an attempt to deduce therefrom the most practicable size of plot and the number of repetitions that would reduce the probable error to dimensions less than the differences to be expected in the trials. But we had no more than arithmetical methods at our command: Dr. Fisher's technique had not been made available for agricultural experimentation. Indeed all this was before the days of Government grants and refined apparatus: we had to improvise our bricks with a minimum of straw. I remember how Mercer constructed a thresher for his little bundles of wheat out of an old bicycle frame, on which he used to mount and pedal away in order to knock out the grain, thus combining research and exercise. Since that time the subject has grown in all directions and a due appreciation of its principles is a necessary part of the equipment of every agricultural experimenter. I would go further and insist that all biological investigation involves a statistical consideration of the results; no organisms exactly repeat one another and a valid conclusion can only be drawn when data are available in sufficient numbers to admit of an estimate of its probability. As is so obvious in human affairs a law of action may only be true statistically and not individually. Hence the importance of the subject with which we are dealing to-day.

https://doi.org/10.23637/ERADOC-1-214

PRINCIPLES OF PLOT EXPERIMENTATION IN RELATION TO THE STATISTICAL INTERPRETATION OF THE RESULTS

By R. A. FISHER

THE greatest source of error in field experimentation is that due to the heterogeneity of the soil. This fact is the conclusion universally to be drawn from uniformity trials, and may be illustrated by the contours of fertility found in a trial carried out by Mercer and Hall at Rothamsted in 1910, in which an acre of wheat, chosen for its apparent uniformity, as is the land allotted to experiments, was treated uniformly and harvested in 500 small plots. The yield, even after smoothing out the variation ascribable to extremely local fluctuations, varied from about 27 bushels per acre in the areas of low fertility to about 37 in the areas of high fertility, a range of about 30 per cent. of the mean yield.

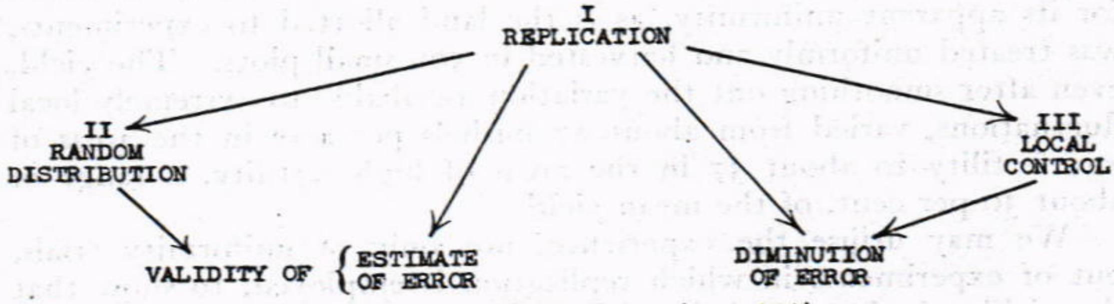
We may utilise the experience, not only of uniformity trials, but of experiments in which replication is employed, to show that the yields of plots of about $^1/_{40}$ acre frequently vary among themselves owing to soil heterogeneity with a standard deviation higher than 10 per cent. of the mean yield. With ordinary care all working errors may be kept down to a much lower figure. It is therefore important to see that the time and labour at our disposal is not wasted in attaining over-meticulous precision in factors which in fact contribute little to the field errors, but that they should be employed where the advantage to be gained is greatest in overcoming the errors due to soil heterogeneity.

In addition to its quantitative importance the uniformity trials have sufficed to establish quite generally (i) that the soil fertility cannot be regarded as distributed at random, but to some extent systematically, so that neighbouring plots are on the average more alike than those further apart, and, on the other hand, (ii) that the distribution is seldom or never so systematic that it could be satisfactorily represented by a simple mathematical formula, as a simple fertility gradient can be represented by a function linear in the co-ordinates.

https://doi.org/10.23637/ERADOC-1-214

The difficulty arising from soil heterogeneity may be overcome in theory by *replication;* we are therefore concerned with the objects which replication is designed to fulfil and with the conditions on which these objects are best achieved. There are two objects, shown in the diagram, which replication is always required to fulfil, namely to diminish the experimental error, and to provide an estimate of the magnitude of those errors.

With respect to the diminution of error by improved replication it should be noted that the precision can in this way be increased indefinitely. It has indeed been argued that since increased replication requires that an experiment must occupy a larger area of land, the soil heterogeneity will thereby be increased, and that in consequence a point will be reached beyond which further replication will give no further increase in accuracy. This seeming difficulty cannot, however, be effective if the different treatments to be compared are always compared locally within relatively small pieces of land, for then only the natural irregularities within such



Principles of Field Experimentation.

small pieces can affect our results. The easiest demonstration of this principle of *Local Control* is to divide the land into a number of blocks, each containing as many plots as there are treatments, of which one is assigned to each treatment. However many blocks may be used in such an experiment the error of our comparisons will be due wholly to soil heterogeneity within blocks, and this element of the heterogeneity has no tendency to increase as the number of blocks is made greater. The increased heterogeneity of the whole area is in fact wholly accounted for by the increasing disparity in yield between different blocks. This element of the soil heterogeneity is, however, entirely eliminated from our comparisons by the arrangement of our experiment. That this fact was not at once realised is due to the use of erroneous methods of estimating the error, which failed to eliminate in the arithmetical procedure elements of variation which had in fact been eliminated from the real errors by the arrangement in the field. This illustrates a point which is of special importance to the question of the estimation of error, namely that it is necessary that our methods of arrangement in the field must be brought rigorously into harmony with the methods of computation to be employed. For given methods of arrangement it is possible that there shall be at most one correct method of computation, and this one we must be able to recognise and to use. For many methods of arrangement, however, no method of estimating the error, which is strictly valid, can possibly exist.

It is thus seen that the second object of replication, the diminution of error, may, if a sufficient number of plots can be used, be carried to any required degree of precision, at least if the primary principle of replication is supplemented by the principle of Local Control. With respect to the first object of replication—to provide an estimate of error—we must now note that, if we are to obtain a strictly valid estimate of error, then it is necessary, in order to satisfy the mathematical conditions on which the use of such an estimate is based, that, apart from such restrictions as are introduced in the complete elimination of certain components of the soil heterogeneity, the different treatments or varieties to be tested shall be arranged at random on the land available. One may say that the heterogeneity of the experimental land is in this way divided into two parts, one of which is totally eliminated from the experiment by the field arrangement, and subsequently in the arithmetical procedure, while the other part is scrupulously randomised in the field arrangement, in order that that portion of it which will be available for the estimation of error shall be truly representative of that portion which necessarily will appear as real errors in our results. The methods by which these principles of experimentation have been worked out in detail are very various, and several examples of these will be given by later speakers.

METHODS OF FIELD EXPERIMENTATION AND THE STATISTICAL ANALYSIS OF THE RESULTS

By JOHN WISHART

The two simplest methods of layout which fulfil the conditions of supplying a valid estimate of error and eliminating a large portion of the soil heterogeneity are (1) the method of Randomised Blocks, and (2) the method of the Latin Square. In what follows these

https://doi.org/10.23637/ERADOC-1-214

methods will be described shortly, with appropriate arithmetical illustrations to indicate how the results of such trials should be analysed.

(1) Method of Randomised Blocks

As an example consider an experiment carried out in 1930 by the Rothamsted staff at the farm of Mr. E. V. Cooke, North Fen, Bourne (Lincs.). The crop was potatoes, and the treatments tested were: superphosphate at the rate of 0, 0·8 and 1·6 cwt. P_2O_5 per acre (approximately 5 and 10 cwt. superphosphate), and sulphate of potash at the rate of 0, 1 and 2 cwt. K_2O per acre (approximately 2 and 4 cwt. sulphate of potash), in all combinations. There were thus nine treatments in all, and these were laid down in four-fold replication. A plan of the experiment is shown below. The area was divided into four equal blocks, each consisting of nine plots, and the nine treatments were allotted at random to the plots within each block. The plot yields, in lb., are given in the table:—

A			B		
0S 2K	1S 2K	1S 1K	0S 1K	0S 2K	1S 1K
2S 1K	0S 1K	2S 0K	2S 0K	1S 0K	2S 1K
1S 0K	2S 2K	0S 0K	1S 2K	2S 2K	0S 0K
1S 1K	0S 2K	2S 2K	1S 0K	2S 0K	0S 2K
0S 0K	1S 0K	1S 2K	1S 1K	2S 2K	1S 2K
2S 1K	2S 0K	0S 1K	0S 0K	0S 1K	2S 1K
C			D		

Area of each plot = $^1/_{70}$ acre.

0S = No super.

1S = 0·8 cwt. P_2O_5 as super.

2S = 1·6 cwt. P_2O_5 as super.

0K = No potash.

1K = 1 cwt. K_2O as Sul./Pot.

2K = 2 cwt. K_2O as Sul./Pot.

Block	*No Potash*			*Single Potash*			*Double Potash*			Total
	0S	1S	2S	0S	1S	2S	0S	1S	2S	
A	372	293	392	360	459	388	344	439	406	3453
B	334	444	437	393	385	434	366	438	439	3670
C	234	291	279	295	339	297	332	413	479	2959
D	262	385	338	335	382	367	297	365	421	3152
Total	1202	1413	1446	1383	1565	1486	1339	1655	1745	13234
							General Mean = 367·61			

The block and treatment totals are given in the margins of the above table, while the general mean of all plots is 367·61. We now regard the deviation from this general mean of each plot yield as being made up in the main of two parts, one equal to the deviation of the block mean from the general mean, and representing the amount by which the yield is influenced by the fertility of the block in which it happens to be situated, and the other equal to the deviation of the mean of the treatment given to that plot from the general mean, and measuring the value of the particular treatment as far as this plot is concerned. The sum of these two portions gives a theoretical value for the plot yield deviation, and the differences, positive and negative, from this theoretical value, of the deviation of the actual plot yield from the general mean of all plots, represent the errors of the experiment, and are used to furnish a standard error for the treatment means, or totals. The sum of squares of these residuals is best obtained by eliminating the contributions due respectively to blocks and treatments from the total sum of squares of deviations of the 36 plot values from the general mean. Instead of subtracting the mean successively from each value, and squaring and adding the remainders, it is only necessary to square and add the actual values as given in the table, and then subtract a correction equal to the square of the grand total divided by 36. The same result could be secured, but with smaller numbers to work with, by subtracting a round number near to the true mean, as, for example, 400, from all the values before squaring. The correction in this case is the square of the new total, divided by 36. The arithmetical working is as follows :—

Sum of squares of 36 plot yields . .	4994120	
Grand total squared and divided by 36 .	4864965·44	
Difference	129154·56	(*a*)
Sum of squares of 9 treatment totals . .	19679370	
Divide by 4 (since each is a total of 4 plots) .	4919842·5	
Subtract as above	4864965·44	
Remainder	54877·06	(*b*)
Sum of squares of 4 block totals . .	44082894	
Divide by 9 (since each is a total of 9 plots) .	4898099·33	
Subtract as above	4864965·44	
	33133·89	(*c*)
(*a*) — (*b*) — (*c*)	41143·61	(*d*)

The sums of squares (*b*), (*c*) and (*d*) must now be divided by the appropriate numbers of degrees of freedom to obtain the corresponding mean squares. There are 35 degrees of freedom in all (one less than the total number of plots), of these there are 8 for treatments (one less than the total number of treatments) and 3 for blocks (one less than the total number of blocks), so that there are 24 remaining for the error part. The results are now set out in the form of an "analysis of variance" table as follows:—

Analysis of Variance due to	*Degrees of Freedom*	*Sums of Squares*	*Mean Square*
Blocks . . .	3	33133·89	11044·63
Treatments . .	8	54877·06	6859·63
Error	24	41143·61	1714·32
Total	35	129154·56	

We note that the method of arrangement has removed, under the heading "blocks," a substantial amount of the soil heterogeneity present. Had the arrangement not been such as to allow for this, the error "mean square" would have been very much higher. The next point to note is that the treatment "mean square" is materially greater than that due to error. With no real treatment differences, these two contributions should have been equal within the limits of sampling error, and we test whether the treatment mean square is significantly greater than the error mean square by finding the ratio of these two quantities, *i.e.* 4·001, and then determining one-half the natural logarithm of the ratio. This gives us the quantity z, which is in this case equal to 0·6932. The same result is obtained by finding the ordinary logarithm of 4·001, *i.e.* 0·6022, and multiplying by 1·1513. From R. A. Fisher's "Statistical Methods for Research Workers," table of z, we find from n_1 (the number of degrees of freedom corresponding to the larger mean square) $= 8$, and n_2 (the number of degrees of freedom corresponding to the smaller mean square) $= 24$, a z value of 0·6064 for the 1 per cent. point. Thus the value 0·6064 would be exceeded by chance only once in a hundred trials, had there been no real treatment differences, and we therefore conclude that as the actual z is 0·6932 the effect of treatment is undoubtedly significant. It remains now to examine the nature of this treatment effect. The square root of the error mean square (1714·32) is 41·40, and this is the estimated standard error of a single plot. The standard error of a total of four plots is got by multiplying this by the square root of 4, and is therefore 82·8, while the standard error of a total of 12 plots is got by multiplying by the square root of 12, and is 143·4. We may therefore rearrange our treatment totals in the following way, in order to bring out the real effects, and we append the appropriate standard errors.

TOTALS OF FOUR PLOTS

	No Super	*Single Super*	*Double Super*	*Total*
No Potash .	1202	1413	1446	4061
Single Potash	1383	1565	1486	4434
Double Potash	1339	1655	1745	4739
Total	3924	4633	4677	13,234

Standard error, 82·8 or 5·63 per cent. ; of margins, 143·4 or 3·25 per cent.

The difference between any two totals should exceed three times the standard error for significance, and an inspection of the table shows that the main effects are a response to superphosphate which is not increased materially by the double dressing, and a response to sulphate of potash which is of such a nature that the yield from the single dressing is intermediate between the other two, although alone the increase is barely significant over the yield of the plots without potash. These results are gleaned from the margins of the above table, which do something to smooth out the irregularities of the individual figures, and which are based on the means of 12 plots. The only additional effect that could emerge would be an interaction between superphosphate and sulphate of potash, *i.e.* where the increase on a plot having both was something more than the sum of the separate increases due to the single nutrients. There is not much sign of such an interaction in the individual figures of the table, but these treatment effects can be tested more precisely, since the experiment is of a balanced type, by breaking up the total sum of squares due to treatment into two parts due respectively to superphosphate and potash, and a third part representing the interaction. The calculations are similar to those already carried through for blocks and treatments, and are as follows :—

Sum of squares of 3 super totals .	58736794
Divide by 12 (since each is a total of 12 plots) . . .	4894732·83
Subtract as in previous work .	4864965·44
	29767·39 (2 degrees of freedom)

Sum of squares of 3 potash totals	58610198	
Divide by 12	4884183·17	
Subtract as above . . .	4864965·44	
	19217·73	(2 degrees of freedom)
Remainder on subtracting from 54877·06	5891·94	(4 degrees of freedom)

The central part of our analysis of variance table may then be given more fully as follows:—

Due to	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	
Super. . .	2	29767·39	14883·70	$z = 1{\cdot}0806$
Potash . .	2	19217·73	9608·86	$z = 0{\cdot}8618$
Interaction .	4	5891·94	1472·98	
[Error . .	24		1714·32]	

z for 1 per cent. point ($n_1 = 2$, $n_2 = 24$) = 0·8626

It is obvious that the interaction is wholly insignificant, the mean square being nearly equal to that for error, while by testing the superphosphate and potash effects by the z test with $n_1 = 2$ and $n_2 = 24$, it appears that both effects are significant. We see therefore that the method of arrangement has not only given us effectively greater replication for the potash and superphosphate comparisons, but has shown that the effects demonstrated for each treatment hold, within the limits of experimental error, over a wide range of the other fertiliser. The analysis is completed by presenting the data of our table in agricultural units, *e.g.* tons per acre. Since the yields given in the body of the table are in lb. per $^4/_{70}$ acre, we must divide them, and the standard error of 82·8, by $^4/_{70} \times 2240$, or 128, while the marginal totals, and the standard error of 143·4, are to be divided by 384 to give marginal means in tons per acre.

(2) Method of the Latin Square

We shall take as an illustration of this method a trial carried out on sugar beet at the South-Eastern Agricultural College, Wye, in 1930. The treatments were (1) control, (2) sulphate of ammonia (3 cwt. per acre) applied with seed, (3) nitrate of soda (equivalent to 3 cwt. sulphate of ammonia) as top dressing, and (4) cyanamide (equivalent to 3 cwt. sulphate of ammonia) applied a few days before sowing. There were four replications of each treatment, and the plots were arranged as in the diagram, so that each treatment occurred once and once only in each row and column of the Latin square. The particular arrangement adopted was one chosen at random out of the total number of Latin squares of this size.

I	S	N	C	O
II.	C	S	O	N
III	O	C	N	S
IV	N	O	S	C

Area of each plot = 1/50 acre.
O = control.
S = sulphate of ammonia.
N = nitrate of soda.
C = cyanamide.

All plots received 4 cwt. superphosphate and 2 cwt. muriate of potash per acre.

YIELD OF UNWASHED ROOTS IN LB.

	Columns				*Total*
Rows	654	661	673	599	2587
	638	621	573	719	2551
	499	581	639	652	2371
	557	477	591	669	2294
Total	2348	2340	2476	2639	9803

General mean, 612·6875

TREATMENT TOTALS

O	S	N	C	*Mean*	*Standard Error*
2148	2518	2576	2561	2450·75	37·92

With this method of arrangement, the rows of plots are replicates of one another, and equally the columns, and a comparison of their totals measures the effect of soil heterogeneity, which can be eliminated in the calculation of the standard error of the treatment means. The deviation from the general mean of the yield of any one plot is in fact regarded as made up in the main of three parts, one equal to the deviation of the row mean from the general mean, another equal to the deviation of the column mean from the general mean, and the third equal to the deviation of the treatment mean from the general mean. The sum of squares of the residuals, or differences of the actual yield deviations from theoretical values made up of these three parts, is then used to furnish an estimate of the required standard error. As in (1) this sum of squares is best obtained by eliminating the contributions due to rows, columns and treatment from the total sum of squares of deviations of the 16 plot values from the general mean. The arithmetical working is as below :—

Sum of squares of 16 plot yields	6068949	
Grand total squared and divided by 16 . .	6006175·56	
Difference	62773·44 (*a*)	(15 degrees of freedom)
Sum of squares of 4 row totals	24084247	
Divide by 4 (since each is a total of 4 plots) . .	6021061·75	
Subtract as above . .	6006175·56	
Remainder	14886·19 (*b*)	(3 degrees of freedom)
Sum of squares of 4 column totals	24083601	
Divide by 4 (since each is a total of 4 plots) . .	6020900·25	
Subtract as above . .	6006175·56	
Remainder	14724·69 (*c*)	(3 degrees of freedom)
Sum of squares of 4 treatment totals	24148725	
Divide by 4 (since each is a total of 4 plots) . .	6037181·25	
Subtract as above . .	6006175·56	
Remainder	31005·69 (*d*)	(3 degrees of freedom)
(*a*) — (*b*) — (*c*) — (*d*) . .	2156·87	(6 degrees of freedom)

Analysis of Variance due to	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	
Rows . .	3	14886·19	4962·06	
Columns .	3	14724·69	4908·23	
Treatments .	3	31005·69	10335·23	$z = 1·6794$
Error . .	6	2156·87	359·48	
Total	15	62773·44		

The significance of the treatment differences is here unquestioned, since the 1 per cent. value of z for $n_1 = 3$ and $n_2 = 6$ is 1·1401, and the value reached, 1·6794, exceeds this. The standard error of a total of four plots is

$$\sqrt{(359{\cdot}48 \times 4)} = 37{\cdot}92, \text{ or } 1{\cdot}55 \text{ per cent. of the mean.}$$

It is clear from this standard error, taken in conjunction with the table of treatment totals given earlier, that the result of the experiment may be summarised by saying that there was a significant increase in yield due to the application of nitrogeneous fertiliser, whatever the form, but that there were no differences in action between the forms of fertiliser, or between the times at which these were applied.

The two experiments selected in this paper are an illustration of the fact that the Latin square method usually provides greater accuracy of comparison (here 1·55 per cent. as against 3·25 per cent. at the lowest). This method is in fact the best where the number of treatments is not too great, *e.g.* up to six or seven, and particularly where all comparisons made are of equal value, as in testing equivalent quantities of the same fertiliser, or in testing different varieties. For experiments involving larger numbers of treatments, as in all cases of the balanced type illustrated in (1), the randomised block method is suitable, but care should be taken that a sufficient degree of replication is provided for.

A brief statement only has been given of the methods of laying out and analysing experimental trials, and each experiment treated usually produces points of its own for consideration, while experiments of a more complex nature than those dealt with here can also be readily carried out. For a more detailed discussion of the principles of the method, and of the method of analysis, the following references should be consulted :—

R. A. Fisher : " Statistical Methods for Research Workers," 3rd edn., 1930, Chapters VII and VIII. Oliver & Boyd. 15s.

R. A. Fisher : " The Arrangement of Field Experiments." *Journal of the Ministry of Agriculture*, Vol. XXXIII, 1926, pp. 503-513.

R. A. Fisher and J. Wishart : " The Arrangement of Field Experiments and the Statistical Reduction of the Results." Imperial Bureau of Soil Science—Technical Communication No. 10. H.M. Stationery Office, 1930. 1s. net.

J. O. Irwin : " Mathematical Theorems involved in the Analysis of Variance." *Journal of the Royal Statistical Society*, Vol. XCIV, 1931, pp. 284-300.

THE TECHNIQUE OF GRASSLAND EXPERIMENTS

By Prof. R. G. STAPLEDON

Plant Breeding Station Aberystwyth

The technique of grassland experiments in twenty minutes! I can only be outrageously eclectic, and I think it will be best, although perhaps egotistical, to confine myself to a discussion of the various methods which we in Wales adopt. It is all a question of the factors we think important and the factors we wish to study.

Personally I think the biotic factor is by far the most important one influencing grassland. I mean the influence of the grazing animal on the sward; this, as far as it is legitimate to use the phrase, is assuredly the *Master Factor*. How can we study this factor? By rigid control of the grazing animal on plots representing different types of grassland. So now I come straight away to my tethered sheep; you can tether bullocks or horses if you have the will and the room. The point about tethering is you can regulate your grazing to any intensity you like and you can do so on small plots—and I shall have a lot more to say about small plots. It's simply a matter of a rule of three sum and a proper system of moving so as not to starve the sheep on the plots which are most intensively grazed. On these bioticas (a biotica is a piece of ground devoted to a study of the biotic factor) of ours our most intensively grazed plots carry the equivalent of 17 sheep per acre per day throughout the grazing season. Let me just incidentally remark, I believe in all field experiments of a research nature we should go at each end far beyond what are deemed by practical men to be economic limits.

A few words on the technique of managing tethered sheep. We use the Scandinavian chains—about 10 feet long, then bifurcating into two lengths of 2 feet each, two sheep being tethered together. The sheep must be moved twice a day—half a length at a time—or they will graze the periphery of their circle to death. They must be given water during appreciable periods of drought—we use ordinary garden saucers. It is desirable that they should be given shelter when tethered in exposed situations, *e.g.* at 1000 feet on the Welsh hills—how to do this cheaply and efficaciously is a problem we are still working at; and I am open to suggestions.

https://doi.org/10.23637/ERADOC-1-214

The plots when a chain wide must go with and not across a hilly field or the sheep will always camp on the higher ground and spoil the plots. For biotic purposes we have trained even two to four-year-old Welsh mountain wethers to tolerate the tether chains—for other tethering experiments to be discussed later we use yearling lambs twelve months old at the commencement of the experiments.

We are investigating six intensities of grazing on five sharply contrasting types of grassland on plots of no larger than about $^1/_{20}$ acre.

So far, you will understand, we are simply using our sheep as controlled defoliating, treading, urinating and excreting machines, and we are not in the least interested in the sheep *qua* sheep.

The control of the sheep represents but one aspect of these ecological investigations, the final results are given by the botanical analysis of the plots—botanical analysis, the bugbear—at least so people think—of all grassland investigations.

Again it is all a matter of what you want, and generally what is wanted is a reliable comparison of the type and degree of such differences as have significance *vis-à-vis* the ultimate welfare of the animals that are turned to graze particular fields.

On physiological—nutritional—and morphological grounds, which I have no time to discuss, the elements which, broadly considered as such, matter in the flora of a sward, and which in the case of many types of investigation it is sufficient to categorise, are (1) clover contribution; (2) miscellaneous weed contribution; (3) contribution of bent and (or) fine-leaved fescues considered as a unit; (4) contribution of Yorkshire fog and (or) soft brome considered as a unit; (5) contribution of that particular species of so-called valuable grass, if any, which may contribute in large amount, say over 10 per cent. to the sward—*i.e.* usually perennial rye-grass, and (6) contribution of other grasses considered as a unit. Unless your treatment, no matter what it be, has substantially influenced the adjustments as between these five or six groups as such—I doubt if it has any economic significance worth considering and still less worth talking about—it is only for quite exceptional purposes that one wants a botanical analysis accurate for all the species as such to the limits of 1 or 2 per cent.—and if you do want it I doubt if it is ever worth the trouble of getting.

There is a tremendous food value difference, thinking in terms of all-the-year-round grazing between the fine-leaved fescue-bent unit and the rye-grass-meadow fescue-cocksfoot unit and a still greater difference between these units and the clover unit, and it is the interplay of these units that, from the botanical point of view, primarily matter and which can be comparatively accurately estimated.

In regard to botanical analysis there are two distinctions we always want to hold clearly before us. Do we want to know what the ration actually offering to the animal is day by day and week by week, or do we want to compare the slow ecological changes which under some series of treatments show themselves on contrasting plots? In the case of my bioticas I am interested only in the ecological changes. For these purposes I am in favour of analyses made by lifting representative turfs and counting the plants. We usually lift ten turfs per plot of $^1/_{100}$ acre—or on larger plots per quadrat of $^1/_{100}$ or $^1/_{500}$ acre. We now lift rectangular turfs of size equal to 6 ins. $\times$ 6 ins. or 1 ft. $\times$ 1 ft.

You can count tillers or you can count plants—if the latter, for each species you have to carry a standard for each of what constitutes a plant—we are out for comparative data—and it does not much matter what your standard is as long as it is rigidly adhered to. It is true that tiller counts favour the multi-tillered species—rough-stalked meadow-grass for instance. But this does not in the least matter—the man or woman who is competent to research on swards is competent to put the proper interpretation on the data so obtained—and as far as I see it that is all that data are for.

Your trained and reliable pasture workers—and you can identify such persons by their personal technique in the matter of taking up a horizontal position on a sward, and in the matter of handling herbage—are competent in appropriate connections to dispense with these laborious counts and instead can estimate *in situ* on the plots to an absolutely sufficient degree of accuracy for a very large number of practical purposes—and practical purposes are our aim. Personally I always marvelled at tea-tasters, wine-tasters and wool-sorters, while the Bobby Jones's and Lindrum's of this world are further proof of the capabilities of the human eye and of human judgment. It is all a matter of what you want your data for and the order of differences you are dealing with.

Now we come to the question of studying the day-by-day interplay of the sward and the animal, grassland research *in excelsis*. Personally I do not like unqualified hay data as an aid to formulating opinions as to grassland management—such data leave the animal out of the question (and all meadows are grazed), while the growth of hay covers only a part of the year—and when precisely is the gate shut on a hay field? A most decisive factor this, as influencing the interplay of species—a piece of information which is far too seldom given in reports on hay experiments; while, incidentally, let me remark the influence of gate shutting can only be investigated critically by resort to folding or tethering.

With studies on the day-by-day interplay between the animal

https://doi.org/10.23637/ERADOC-1-214

and his herbage it is again a question of control of the animal and of botanical analysis. Obviously you must control the animal, because you cannot take reliable samples of what the animal is eating or to give you pasture yield on fields ordinarily extensively grazed. True, you can put wire mesh cages over quadrat areas and sample these on any basis that seems good to you. This method has been employed in America—we have tried it but think our method of controlling the animal preferable. In brief our method is to sample in front of the controlled grazing animal.

The animal can be controlled by tethering and this is the method we are now using to an ever-increasing extent.

At first we used and are still to some extent using, penned plots in size varying from $^1/_{100}$ to $^1/_{50}$ acre—plots of the order of $^1/_{10}$ acre only being occasionally used. When using plots one to four sheep are introduced according to the size of the plot and the amount of eatables offering—24 to 72 hours completes the grazing. Whether we use plots or tether we usually graze and therefore sample on a three weeks' rotation.

To obtain yield data we sample within a rectangular mesh equivalent to a square foot on small plots, and a square yard on larger plots. We take at each grazing ten samples per $^1/_{100}$ acre plot—or in the case of more extensive work, five samples to a quadrat of $^1/_{500}$ acre. We cut hard to ground level with sheep shears within the mesh. You must sample by this means after each grazing—this gives what the animals have left—and sample again before introducing the animals for the subsequent grazing. Thus for example the weight given by the post-grazing sample taken on 31st March, subtracted from the weight given by the pre-grazing sample on 21st April will give you the yield of herbage produced during the first three weeks of April. Note, firstly, you cut as low as is possible —lower than the animals graze. Note, secondly, you are sampling, and therefore you are not subjecting your whole plot at each grazing to drastic mowing machine-wise defoliation; and *note thirdly*, you are causing your plots to be trodden, urinated and excreted upon. After the post-grazing sample we spread the droppings.

Ideally it would be nice to give your yield data as oven-dry fodder, but owing to limitation of time and facilities we only do this in the case of very special experiments. Normally we make what is tantamount to hay in scrim bags and give our data as air-dry fodder. We weigh in a fine dry day after a run of five dry days.

This is the point—grassland problems are essentially complex problems, and in order to enunciate helpful dicta relative to the economic management of grassland you have got to understand the interplay of innumerable factors. You have therefore to deal with

https://doi.org/10.23637/ERADOC-1-214

a prodigious number of plots and of samples and you are usually dealing with differences of considerable magnitude. The only thing that matters is whether your data ellucidate your problem? If the answer is in the affirmative your technique is meeting its case.

Now for botanical analysis again, and this time the analysis of the ration the animal is eating. We quarter down our samples to 1 lb. During the commencement of these researches we actually separated all these herbage samples—and they have run into thousands—into the appropriate categories; and it is literally marvellous what well-trained girl assistants can do in this respect even with short pasture herbage.

Note in passing, please, that not even the best trained girl can analyse the mince-meat type of herbage delivered by a garden mowing machine. Now that we have all served a long apprenticeship—we of the research staff and the senior girls are largely adopting Mr. William Davies' admirable scheme of estimating. You divide your little 1 lb. heap of herbage into ten approximately equal heaps. You give ten marks to each heap. Then for each heap allocate out the ten marks between your categories—noting "trace" against the category represented but not worthy of a mark. Adopting this means a properly trained person will get a percentage by weight contribution for the categories as near as quite literally does not matter to that obtained by all the laborious separations and weighings. The people concerned have tested themselves repeatedly. It is not too much to say that when we had the courage to adopt the Estimating Method as part of the regular technique employed at the Station, by the stroke of a pen, as it were, we were able to give a tremendous impetus to all our endeavours, and I think we are now definitely at grips with the grassland problem. We have, in short, developed a technique based on a method of controlling the animal, a method of sampling and a method of obtaining botanical data on a very large scale which is making it increasingly possible for us to study the interplay of the innumerable factors involved.

A word as to live weight increase. We are interested in this because we want to test our pedigree strains and because it is an important aspect of the grassland problem.

To get the best results I am certain you want rigid control of your animals. You may either rotationally graze over small folds, or you may tether—both methods are infinitely superior to merely dividing off large plots of a sufficient size to carry a sufficient number of experimental animals as one extensive unit. Our experience, and it now goes over three seasons—but mark you, I am only talking in terms of sheep—is overwhelmingly in favour of tethering. Such is our faith in tethering that we are setting up an experiment this year to test rye-grass against cocksfoot that will entail the use of forty-eight pairs of tethered sheep. Here are the advantages of tethering: (1) you can replicate your plots sufficiently well, I should imagine, to satisfy Dr. Fisher himself; (2) your experimental sheep will be handled and examined twice a day—they will be properly looked after; (3) your grazing will be uniform—soil inequalities will therefore count less; (4) your carrying capacity of thriving animals will be increased per unit of area, and therefore your number of experimental animals per unit of area can be greater than under more extensive methods. Statistically what a blessing this is to those of us who have to operate on a limited area of ground, or who are dealing with pedigree strains of grasses of which seed is necessarily limited, or for that matter to all who appreciate the multiple factor aspect of grassland problems.

I would like, in conclusion, to say three words about small plots. I have said the grassland problem is a complex problem; it is largely an ecological problem, and although eventually no doubt we shall be able to pack an increasing number of its sub-problems into the laboratory and the greenhouse, it will always remain in its wider and more definitely economic aspects a problem that will lend itself to a marked degree to that type of mind which has about it a considerable streak of the naturalist—a turn of mind which has an outdoor rather than a laboratory way of looking at things. I like the small plot not only because it is an adjunct to adequate replication for statistical purposes, but because it makes you look at it as a whole—makes you concentrate attention upon it—and thus the greater your number of replications by that much the deeper and more intensive your contemplations. In grassland studies you must collect data, and you want data of a degree and precision adequate to your term of reference; but I believe the problems are solved just as much—indeed I am heterodox enough to say perhaps more—by what one notices in the field when collecting the data as by the actual yields or other precise—but usually end-stage—results which the analysis of your data gives you.

If I may detain you one more minute, I would like to say a word as to control plots and a word as to lay-out. In seeds mixture work the ideal control—and a most valuable plot—is your "no seeding" plot, and such a plot should, I think, always be generously introduced if necessary as an extra control. For one thing it is likely to tell you from where a bit of your wild white clover has come from, and it is the basal plot relative to competition—weeds have so much to do with it. As to lay-out I would like to emphasise the enormous amount that is to be learned by a "one issue" experiment carried over the widest possible area and the greatest possible diversity of conditions. For this purpose I like a strip, say a harrow or two harrow width, carried right through the middle of a field and if possible on and over contiguous fields. Two controls and your treatment—up hill down dale, over every sort of soil. We have quite a number of these, our largest being nearly a mile long, in connection with our work on the open hills. What, for example, would not a basic slag strip right across England have taught us, and I would like to join up Rothamsted and Aberystwyth with a single mixture strip.

THE TECHNIQUE OF VARIETY TRIALS

By S. F. ARMSTRONG

National Institute of Agricultural Botany

Introduction

This is rapidly becoming an extensive subject, and as our programme to-day is lengthy I shall confine myself to such points as I feel are especially worthy of emphasis. These points are largely the outcome of experience gained by the staff of the National Institute of Agricultural Botany while engaged in actual field trials.

The first essential for those engaged in this work is that they should realise something of the difficulties and complexities of their task. We are dealing with living organisms. We are setting out to get a measure of their output—*yield*—itself the result of a large variety of causes. We are attempting to measure that even more elusive and undefinable thing—*quality*; and besides this we have still to take into account many other " crop behaviour points " which often to a large degree modify the " crop value " in farming practice. Moreover, we are dealing with " variety trials," that is to say, with the comparison of closely related plants. Therefore the differences we are attempting to measure are not usually of a large or obvious kind. For these and other reasons a well-defined and scientific technique is essential to success.

In variety trials it is necessary that we have a suitable standard of measurement by which we may judge the comparative performance of any variety as regards yield, quality or value of produce. The best known means available to us is to place a similar organism (or mass of organisms) under similar conditions. We employ a closely related variety as a standard or control and in this way obtain a

https://doi.org/10.23637/ERADOC-1-214

relative measure of the performance of the test variety. From beginning to end of a variety trial it is therefore essential that the treatment of the control and the test variety shall be identical.

It will be obvious that a given variety cannot be used as a suitable standard of comparison against any and every variety of the same crop. To be suitable a control should mature at nearly the same time as the varieties tested alongside it. Also it should be a fairly widely grown variety of good standing either in general or local agricultural practice.

I would emphasise a further point. If the information we derive from variety trials is to be of real value in agricultural practice such trials must be carried out in the normal manner of cultivation and under ordinary farming conditions.

Choice of Land

This requires due care. Some unevenness is sure to exist, but land should be avoided for variety trials where gross irregularities occur either in depth, slope, water supply, etc. The more level and uniform the soil the more likely it is that any differences obtained between two varieties are real varietal differences and not due to soil variation, *i.e.* the results will possess a higher statistical significance.

It is worth while enquiring into the performance of previous crops to get an approximate measure of the uniformity and fertility of the land. Such knowledge will serve as a guide to the rational treatment of the experimental crops. Indeed, all possible available information about the soil, its previous cropping, manuring, etc., should be gathered before laying down trials. It should be ascertained beforehand that game or vermin do not exist in such numbers as to be a real danger to crops. The land should be ample in area so that plots may be at least 20 yards clear of hedges, shade of trees, etc.

Lay-out of Trials

On the respective merits of different lay-outs I do not feel qualified to speak. I shall therefore confine myself to such points as arise in the proper management of the lay-outs which are at present adopted by the National Institute of Agricultural Botany. It is probable that the perfect plot arrangement for variety trials is unworkable in ordinary practice, and therefore some form of compromise between the two aspects of the matter must be adopted.

In the Institute's cereal trials ten pairs of comparative strips are laid down for each separate trial. This number is found to be ample, whilst it does not involve any serious practical difficulty. The strips are sown at right angles to the direction of previously formed furrows and ridges. Dr. Beaven's half-drill-strip method is employed for all cereals, a wide 16 coulter steerage drill being used. The control and variety strips are thus sown side by side in a balanced manner which makes direct comparison simple. As "interference" is likely to occur between different varieties of cereals a so-called "interference strip" is interpolated between each variety and control strip. Also as irregular spacing would frequently occur at the wheel track due to overlap or underlap another "interference strip" is arranged for at this point. These "interference strips" are discarded at harvest.

A complete trial occupies half an acre of land, plus margins. Each comparative strip for weighing is normally $^1/_{80}$ of an acre. There are ten $^1/_{80}$ acre strips of the test variety for comparison with the same number of control strips. The rest are discard strips.

It is most important that the spacing of the coulters shall be accurate and equal on both sides of the drill. Each comparative strip must be of equal width, otherwise a serious constant error is introduced. A flat wooden rod (1 inch by 4 inches), with notches cut to indicate the correct coulter positions, is much better for spacing purposes than a tape measure. This should be used at the beginning and end of every trial. If any serious "shift" does occur it may be necessary to measure across the rows at brairding time, or across the stubbles after harvest, to discover what the average spacing actually is.

In the case of root crops, a single row is considered as a strip. With such crops it is possible to place the ten rows in a "scattered" layout over the entire trial area. The rows are all first correctly marked out with an empty drill; each row is then numbered and the strains sown in their respective rows by means of a one-row hand drill. One variety is selected as control. The mean of all the yields is used for the calculation of the Probable Error and each strain can be directly compared with another. Normally no interference strips are used where only strains of a similar type are grown in one set of trials. If, however, it is desired to test a large topped strain against a small topped one interference rows are introduced, *e.g.* in some sugar beet trials.

Grasses and clovers are sown on a scattered lay-out with intervening interference strips. For these a seed barrow is used, half the seed being sown in one direction and half in the opposite direction to ensure greater uniformity.

Each drill strip (or row) receives a distinctive number by which it can be referred to if special observations are necessary. We dispense with the ordinary wooden label as far as possible until all cultivations have been completed. After this the varieties and controls are indicated by suitable labels. In all cases, however, carefully drawn plans are placed in the field notebook and we depend on these plans rather than on labels to show the position of the strips.

When plots are sown by hand (*e.g.* small checker board plots) it is possible to define the exact limits of each plot. In laying down field variety trials it is not practicable to do this, or even desirable. The ends of strips are always more or less uneven and ragged owing to irregular starting of the drill, to inspections, etc. All strips are therefore drilled several yards longer than they are actually required to be. These extra lengths at the ends are useful for inspection, etc., and also serve to protect the trial area proper. At harvest these surplus lengths are removed, leaving the actual trial area isolated ready for cutting.

Plans for crop protection form a very important part of the work in variety trials. Indecisive results and even complete failures are frequently due to pests of various kinds. No pest is more serious in cereal trails than the common sparrow. It is not an easy matter to keep these birds entirely off the crop, even though a man is constantly on the spot. Clock guns are helpful. This year we are attempting to meet the difficulty by an additional precaution. We are surrounding each block of trials with a strip of an earlier ripening variety in the hope that this may attract the attention of the birds from the actual trial.

Rate of Seeding

A full and regular plant should be aimed at. With root crops extra heavy seedings are used, *e.g.* 12 to 14 lb. of mangold seed. The plants can be thinned and spaced to any desired distance.

With cereals the problem of seeding is different, as there cannot be excess seeding and thinning out afterwards. The question then arises—Is it more correct to sow the control and test variety seed in equal numbers or in equal weights per unit area? The Institute normally follows the method of sowing equal weights for the reason that in agricultural practice it is the general custom to sow a given volume or weight of seed per acre. Different varieties of cereal grains vary considerably in size and shape—expecially in oats. Therefore to even out the seed rate of two varieties (either by number or weight of seed sown) necessitates a special form of drill if the two varieties are to be sown by the same implement at the same time. No English make of drill will do this. The Institute uses a Czecho-Slovakian type known as "The Melichar" in which the cups are expansible. The two halves of the drill are specially made to work independently of each other. The cups on either side may be increased or reduced in size, so that seeds of widely different sizes may be sown in equal weights or numbers at the same time.

It is of course necessary to make preliminary tests to discover the correct setting of the cups. In making such tests it is important to be sure not only that equal weights of seed are delivered on either side, but also that all the sprouts are fed uniformly throughout.

Seed Supply

As is well known the size, weight, moisture content, purity and germination of grain are all influenced by the soil and climatic conditions of the year of growth, and these points are still further modified by the condition of harvesting, threshing, cleaning and storage which they undergo. For obtaining accurate comparisons it is therefore essential that the seed of variety and control shall have been grown under similar conditions. Both stocks should be grown in the same year in the same locality, and if possible side by side. This is the invariable aim of the National Institute of Agricultural Botany and it also takes steps to see that the conditions of handling and storage are identical from the time the crops are harvested until the grain reaches its destination for sowing.

Cultivations

In order that our results may apply to actual farm practice we follow what are considered to be the most approved practices of the district in which the trial is situated. This applies to the preparation of the soil and seed bed, the time of sowing, the seed rate, distance apart of rows, etc. The same is true of manuring and all subsequent cultivations.

The essential point insisted upon is that throughout, the variety under test shall receive precisely the same treatment as its control, and also at the same time. Thus, such operations as rolling, application of fertilisers, etc., are carried out at right angles to the drill strips. In setting out or singling roots the work is done in definite strips across all the varieties. In this way any difference in type of work is spread over all strains alike, and also each receives attention alike in point of time. This is a very important matter in variety trials.

Harvesting

For any crop we first decide upon a base line. This is fixed quite clear from the ragged ends of strips. From the base line the length of strip to be cut and weighed is measured off.

In the case of cereals the strips are cut one at a time, all in the same direction. Either a reaping machine or a binder is used for this purpose. If the crop is lodged it may be necessary for men to lift up the crop on each strip previous to cutting. It is essential that the whole of each strip be cut evenly, and that no portion of a neighbouring strip be taken in. If odd straws are missed by the reaper these are cut off by hand and placed in the strip to which they belong. As each strip is cut the sheaves are tied with string of a distinctive colour to indicate variety, while the strip itself is labelled with its proper number.

As soon as possible after cutting, each strip is weighed, and all the control produce placed in one stack and the variety in another. If storms or other circumstances put a stop to operations when either cutting, stooking or weighing is in progress the work is always left off at a point where an equal number of comparative strips is completed.

Strip Weighing, Threshing

Spring balances are convenient for taking weights of the strip produce. As they are subject to variation they should be tested frequently. Weighing cannot be done in a satisfactory manner when wind is strong, or in wet weather. The half-drill strip lay-out has, however, the great advantage that even if weather conditions are rather adverse the method of weighing the strips in comparative pairs gives comparable results. In all cases, as a precaution, the canvas weighing slings are weighed every half hour, or even more frequently, to find if any change in weight is taking place.

Roots are conveniently weighed in baskets holding about 100 lb. Wire baskets are far superior to wicker baskets. The latter soon collect mud and absorb much water. Roots like sugar beet, to which soil adheres, must be thoroughly washed if yield differences are to be correctly estimated.

In obtaining weights of cereal grains the thresher employed should be of a simple type which may quickly be cleaned out. The Institute uses a small thresher manufactured by Messrs. Garvie & Sons of Aberdeen. It has a 2-feet drum, and is without screen or elevator. Separation of the grain from chaff and debris is effected by the use of a fan in conjunction with two sieves. The object is to secure *all* the grain from the straw with the minimum of breakage. If the grain is much cut or broken estimations of quality are rendered difficult or impossible.

Sampling of Produce

To compare *yields* accurately the moisture content of the produce must be known. Samples for "moisture content" are taken at the same time as the grain weights and placed in sealed bottles.

Quality is taken into account as well as yield. In the case of wheat and barley sufficiently large samples are drawn for milling, baking or malting tests. These tests are conducted by experts under controlled conditions. In the case of oats the percentage of husk is determined. Samples are also drawn for local market valuations to be made by competent millers or grain merchants.

In the case of roots the percentages of dry matter and sugar are determined. When sampling sugar beet whole roots are taken; other roots are cored. Two duplicate samples, each of fifty roots (or cores) are drawn of each strain. Each of the ten rows contributes ten roots to the samples. End roots are excluded, and every *n*th root is taken in such a manner that the fifty roots of sample A are drawn from every portion of the trial area, and so also for sample B. Cores are wrapped in water-proof paper and packed in tins, and all samples are despatched by passenger train the same day to the laboratory.

Recording Observations

It is most important that accurate records of crop behaviour be kept at all stages of growth. Such records may indeed afford the clue to exceptionally good or bad results. Where two varieties are apparently equal in yield and quality the possession by one of them of an excellent additional feature may turn the scale entirely in its favour. In making such records we judge constantly by comparison with the control. All statements are relative to this standard. It is therefore essential that all observations be committed to writing *on the spot*. Mental impressions are fleeting things and can only be accurately conveyed by writing them down whilst they are clear and vivid in the mind.

One has constantly to guard against the use of loose or vague expressions. For example, the statement that "a certain amount of disease is present" gives no definite or comparative information at all. The following suggestions are made: where counts or measurements are possible these should be taken, *e.g.* number of plants

https://doi.org/10.23637/ERADOC-1-214

missing in a row; number of "bolters" per row; length of straw in inches.

When word description is difficult, or lengthy, a simple sketch may serve far better, *e.g.* to indicate the shape of a mangold or beet-root or its relative position in the ground.

When counts or measurements are out of the question a scale of points may be usefully employed, *e.g.* to denote severity of rust attack, amount of lodging, etc.

Care is necessary if weights are to be recorded accurately. The produce of a strip may require two or several separate weighings. Each of these weighings is recorded so that the totals may be independently checked. Each weight is given in lb. and includes the basket or weighing sling. The tare weights are deducted afterwards, and not at the time of weighing.

Sheaves of cereals and grain are weighed to the nearest half lb.; straw and roots to the nearest lb. In recording fractional weights in decimals less than ·5 is expressed as the nearest whole number below it, and ·5 or any decimal greater than this is expressed as the next whole number above it.

Statement of Results

For the purpose of direct comparison the yield of a variety is most conveniently given as a percentage of the control. These percentages are based on dry weight yields. The size of the calculated probable error then indicates whether any difference may be considered as significant or not.

From the farmers' standpoint it is also desirable that results of variety trials should be stated on a comparative financial basis. In any such statement not only yield and quality have to be considered, but also any factors that appreciably affect the cost of growing, or marketing, *e.g.* in sugar beet the size of top, bolting, fangs, ease of lifting, etc.

Conclusion

Finally there must be *continuity of work* at any given centre. Variety trials are greatly influenced by soil, climate and season. Single year's results are of little value by themselves. It is only after the accumulation of at least three years' results (and often longer periods) that information can be published with any real confidence.

Also in variety trials adequate supervision is essential if all the necessary details of technique are to be properly attended to. Thoroughness in a few trials is far more likely to produce helpful

https://doi.org/10.23637/ERADOC-1-214

information than numerous trials conducted under rush conditions. Statistical methods may serve to interpret results, but they are no remedy for agricultural errors. And great patience, perseverance and constant vigilance are prime requisites of workers in this particular branch of investigation.

MULTIPLE SCHEMES OF FIELD EXPERIMENTS

By A. H. LEWIS, A. D. MANSON and J. PROCTER

Imperial Chemical Industries

Introduction

JUST as repetition of plots of the same treatment in a single experiment adds to our information in that it enables more precise comparisons to be made between the treatments tested, so repetition of experiments is of value in substantiating conclusions reached for a single experiment and in testing the treatment over a variety of soil and climatic conditions and over more than one season. It is often found, for example, that the seasonal differences may be of a quite different order of magnitude from the differences between treatments which are often relatively small, while, what is more important, it may be found that the response to the treatments under test cannot be easily predicted. In some seasons a treatment may show a response; in others it may be ineffective or may even depress the yield. Similarly, variation in response may be found with variation in kind of soil or climate.

By multiple schemes of experiments is meant merely a series of experiments of the same type conducted at a number of different centres. The individual experiments should all be alike in that they should contain the same number of treatments at the same rate of application. They should also be of the same form of experimental design, *i.e.* they should all be randomized blocks or Latin squares with the same number of replicates. Although of the same form of experimental design, the random lay-out should not be standardised. It is essential that the lay-out at each centre should be selected at random from all the possible lay-outs of the same form of design. When a scheme of experiments satisfies the above conditions the data obtained can be analysed as a composite whole. The average effects of the treatments, the difference between centres and the interaction of treatment and centre and season can be estimated.

It is quite feasible to conduct a number of experiments of a fairly simple character at a number of centres, and at the base or research station, where facilities are better, to test the same treatments on a more precise basis, *e.g.* four treatments might be tested in duplicate or triplicate blocks at the outside centres but at the base a more elaborate experiment could be conducted.

1930 *Experiments*

During the 1930 season Imperial Chemical Industries Ltd. carried out a number of multiple schemes, both in the British Isles and in several foreign countries, including Portugal, Egypt and China.

The schemes started from a small beginning. Experience gained during 1927-29 showed that absolute control by the experimenter at every important stage of the experiment was necessary. It was also found that randomised experiments involving a number of replicates of *small* plots of about $^1/_{40}$ acre in area were more valuable for demonstration purposes than large plots of $^1/_2$ to 1 acre, the Latin square form of layout being particularly suitable for this purpose.

Thus the use of small plots increased the accuracy of the experiments and lowered their cost whilst increasing their demonstrational value. Another great advantage in using small plots is that co-operation with the farmer is facilitated. Large plots mean inconvenience and loss of time to the farmer as his machinery and help are required to a considerable extent, whilst, with small plots, these difficulties are reduced to a minimum.

The success of multiple schemes, carried out over a large area, depended upon centralised control, detailed organisation and meticulous care on the part of the experimenter, and last, but not least, the goodwill of the farmer. Co-operation of farmers was easily secured when the nature of the experiments was explained, and when it was made clear that they would be absolved, where necessary, from all work in the conduct of the experiments, and that we were always prepared to hire casual labour.

It is essential that the experiments be carried out by men with a thorough training in the practical side of field experimental work.

Experiments in British Isles

The experiments carried out on a number of crops in the British Isles in 1930 involved two fertiliser treatments and a control, so that the 3×3 Latin square form of lay-out was adopted. Although a single 3 × 3 Latin square is of relatively little value, since no great reliability can be placed in the standard error obtained and, in any case, differences have to be very large for significance, the results of a number of 3 × 3 Latin squares can be grouped together and treated statistically as already explained.

Labour

When the experimenter was assisted by one casual labourer, the two operations of laying down the experiment and application of fertilisers could be completed on a cereal, hay or root crop in about three hours. A cereal or hay trial was easily harvested in one day by two casual labourers. Four casual labourers could lift a trial on roots in one day. Grading and weighing of potatoes was carried out on the second day.

Cost

The cost of all casual labour on any root crop never exceeded £3 to £4, whilst hay and cereal trials could be done for at least half that amount. It frequently occurred, especially with the hay and cereal trials, that the farmer offered the casual labour and would not consider any compensation.

Threshing Cereals

It was necessary to undertake the threshing of cereal trials ourselves, as it was not considered satisfactory to thresh each small plot separately with the ordinary farm thresher. A portable thresher with an 18-inch drum mounted on a commercial motor, the thresher being driven by a 4 h.p. engine, was used for this purpose. The total cost was about £400. Using this machine, the threshing and weighing of the produce from nine plots was easily completed in four hours. That the outfit gave complete satisfaction will be realised from the fact that twenty-four trials distributed throughout England, Scotland and Ireland were threshed in a month. During the threshing itinerary in England and the Irish Free State one trial per day was averaged as the weather was good.

Degree of Accuracy

The figures given in Table I give an idea of the standard of accuracy obtained. It will be seen that all the standard errors are satisfactory :—

TABLE I

Country	Crop	No. of Centres	Layout. Treatment × Replicates	Standard Error. Percentage of Mean: Average Results for all Centres	Standard Error. Percentage of Mean: Average Results for Individual Centres
BRITISH ISLES	Hay	16	3 × 3	0·52	1·93
		7	3 × 3	3·09	2·18
		3	3 × 3	3·12	3·29
	Oats Grain	13	3 × 3	1·84	3·7
	Oats Straw	13	3 × 3	1·69	5·03
	Potatoes	14	3 × 3	2·57	4·62
	Sugar beet Washed beet	13	3 × 3	2·26	3·59
	Sugar beet Tops	13	3 × 3	2·54	5·10
PORTUGAL	Maize Grain	5	4 × 4	1·51	—
	Maize Straw	5	4 × 4	2·25	—
	Maize Grain	4	4 × 4	2·26	—
	Maize Grain	4	4 × 4	3·68	—
	Potatoes	7	4 × 4	2·63	—
CANADA	Potatoes	31	2 × 1	1·7	—
		21	2 × 1	2·1	—
		10	3 × 1	2·8	—

Value of Multiple Experiments

To show the value of multiple schemes we may take an example. The example demonstrates the effect of rainfall on the response of maize to fertilisers and is taken from the results of an experiment conducted in Portugal in 1930. In Fig. 1 the distribution of the centres is shown and in Fig. 2 the results of fertilization are represented diagrammatically. It will be seen that a response to nitrogen, as reflected in the yields of grain and straw, was obtained in the wet areas whilst nitrogen had a depressing effect on yield in the dry areas.

Conclusion

In conclusion a plea should be made for an improvement in the general standard of field experiments. Although a considerable improvement has been effected in recent years due to the adoption of

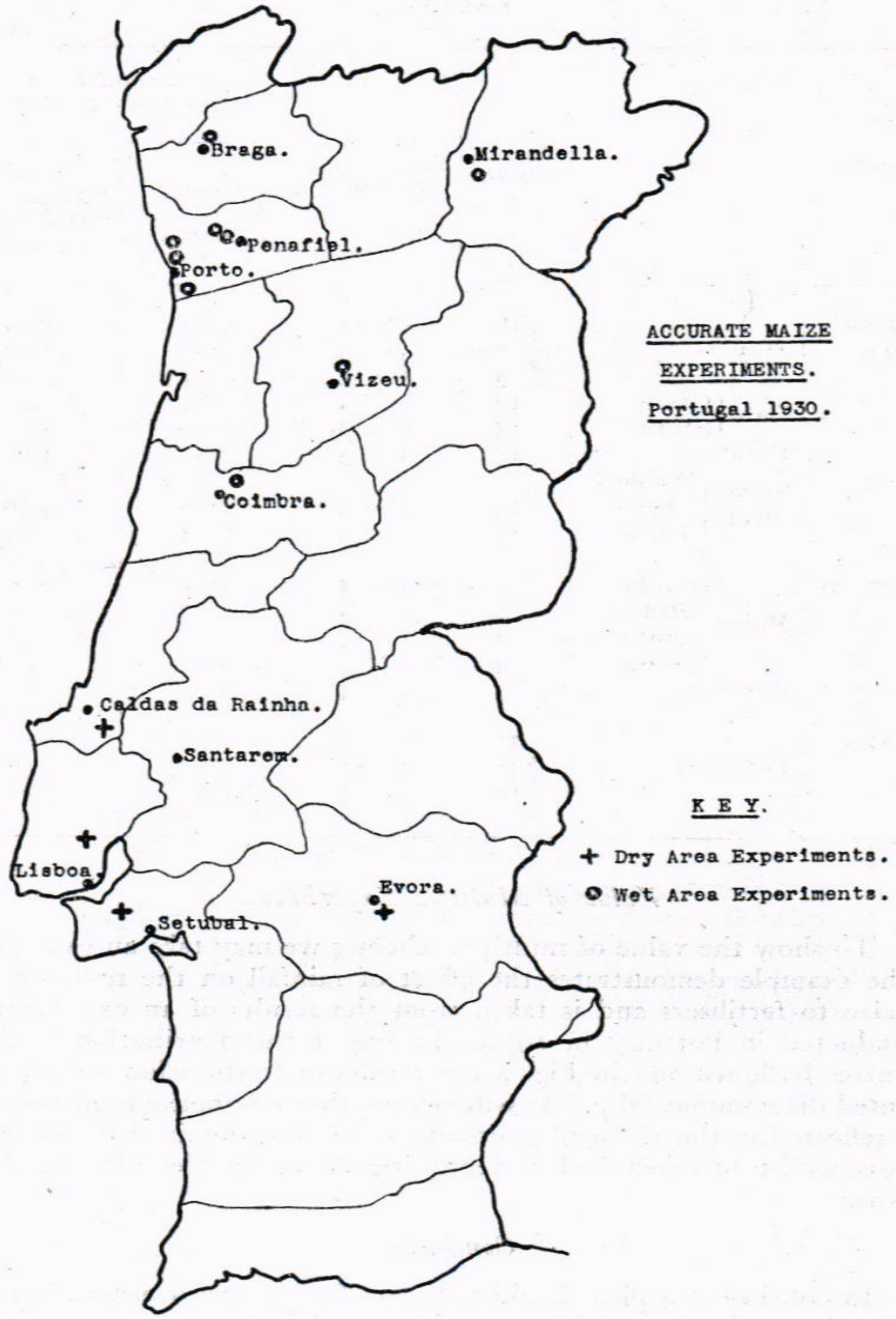

Braga.
Mirandella.
Penafiel.
Porto.
ACCURATE MAIZE
EXPERIMENTS.
Portugal 1930.
Vizeu.
Coimbra.
Caldas da Rainha.
Santarem.
K E Y.
+ Dry Area Experiments.
Wet Area Experiments.
Lisboa.
Evora.
Setubal.

FIG. 1.

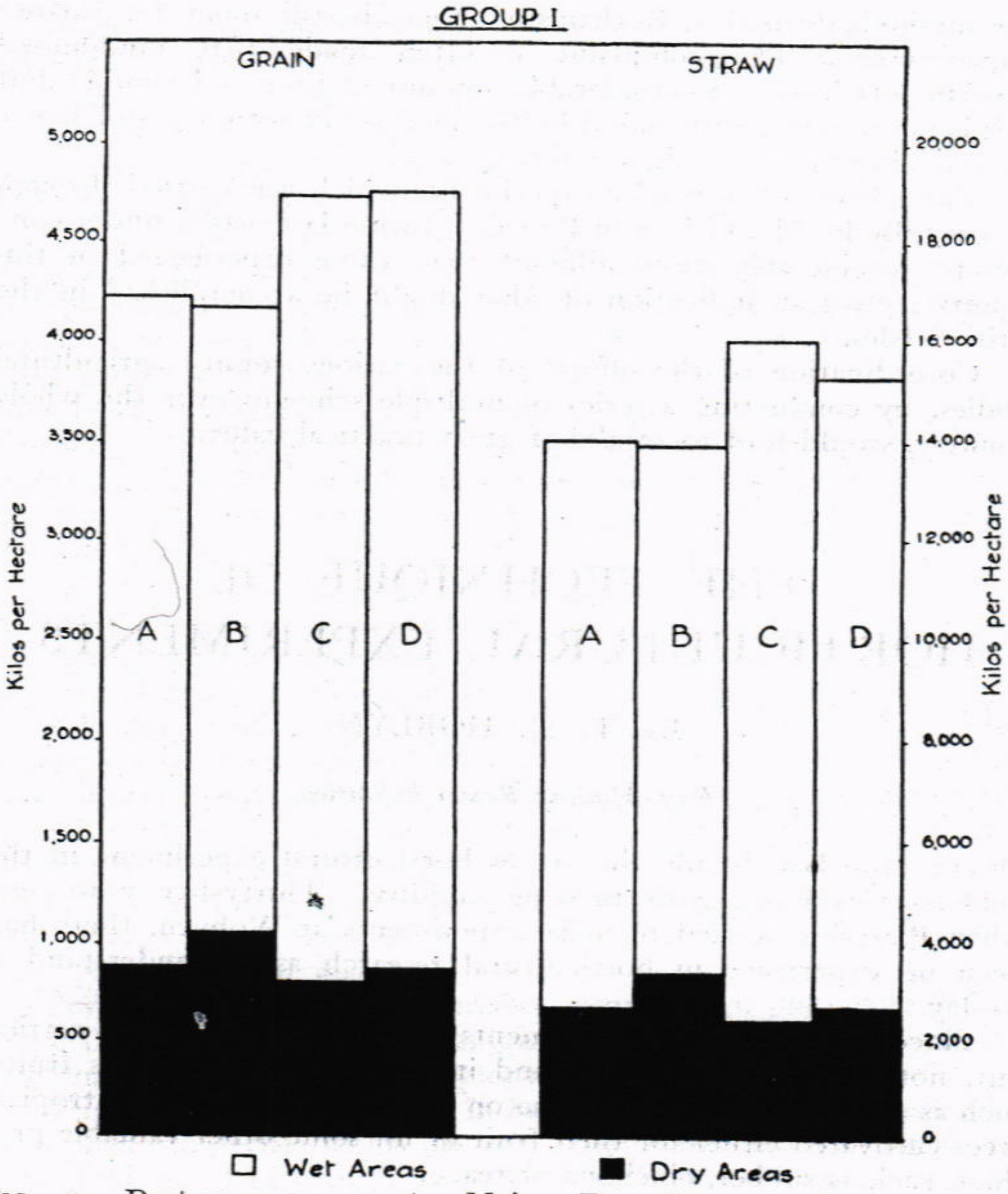


FIG. 2.—Portuguese accurate Maize Experiments. Season 1930. Graphs showing effect of Treatments on Yield in Wet and Dry Areas.

TREATMENTS

A No Fertilisers
B 700 kilos 12 per cent. Superphosphate
100 ,, Sulphate of Potash
C 200 ,, Sulphate of Ammonia
100 ,, Sulphate of Potash
D 200 ,, Sulphate of Ammonia
700 ,, 12 per cent. Superphosphate
100 ,, Sulphate of Potash

} Per Hectare

the methods devised at Rothamsted, there is still room for further improvement. The complaint is often made that randomised experiments involve a considerable amount of time and trouble, but this is more than compensated by the increase in accuracy and hence in reliability of the results obtained.

The scheme of over fifty experiments which was carried through successfully by Mr. Grieve in Portugal during last season under conditions considerably more difficult than those experienced in this country, gives an indication of what might be accomplished in the British Isles.

Co-ordination of the efforts of the various county agricultural bodies, by conducting a series of multiple schemes over the whole country, would lead to results of great practical value.

THE TECHNIQUE OF HORTICULTURAL EXPERIMENTS

By T. N. HOBLYN

East Malling Research Station

During the last decade the art of horticultural experiment in the field has developed with amazing rapidity. Thirty-five years ago, when Pickering started to make experiments at Woburn, there had been no experience in horticultural research as we understand it to-day.

Since that time many experiments on fruit trees have been carried out, not only in this country and in America on deciduous fruits, such as applies and pears, but also on many tropical and sub-tropical trees cultivated either for their fruit or for some other valuable product, such as rubber, cinchona or tea.

The early experimenters soon found that they were up against problems of lay-out and technique which presented many complications not usually associated with agricultural experiments, and in consequence a system of experimentation has been gradually built up which is in some respects radically different from that used in agriculture.

The first and obvious difference is in the nature and longevity of the plant. Practically all horticultural plants are perennials; and thus an experiment once planned and planted must stay *in situ* often for ten or fifteen years before any results begin to appear.

Such an experiment involves the recording of the tree's history for perhaps thirty or forty years, a period when all kinds of seasons will be experienced and during which individual trees are liable to damage or accident from any number of sources, all of which must be protected against as far as possible, since such an accident as, say, an aphis attack may alter the performance of a tree for several years.

The second fundamental difference, which is perhaps not so obvious to the layman, is due to the fact that practically none of the plants grown in horticulture will repeat themselves exactly if raised from seed. Methods of vegetative propagation have, it is true, made possible the standardisation of many kinds of fruit tree; thus black currants are produced from cuttings, and apple trees by budding or grafting a scion from the original parent tree on to a rootstock which has been obtained by layering or stooling.

There are, however, many crops, *e.g.* coconuts and other monocotyledonous plants, which cannot be reproduced by other than seminal methods. In other cases, *e.g.* citrus fruits, or cacao, while it may be possible to reproduce a given variety by budding or grafting, the necessary rootstock cannot easily be induced to root by vegetative means and has to be produced from seed.

All this seedling material exhibits immense variability. Thus an apple tree on one rootstock may be ten times the size of an otherwise similar tree on another. Seedling cacao again, as usually grown, is said to "exhibit a range of variability covering all the existing horticultural varieties of apples."

It is therefore essential in nearly all horticultural experiments (and always where trees are concerned) that observations should be taken on individual plants.

For example, in a manurial trial on, say, cacao, if plots of trees were used, the effect of the manure might be different upon each tree and the sum effect upon the whole plot nothing at all.

The most urgent problem in horticulture is thus the production of uniform material. This has to a large extent been accomplished for deciduous fruits, and research work in vegetative propagation and in breeding is in progress already for a great many tropical and sub-tropical crops. However, even when this is accomplished, owing to the manifold accidents which may occur in the long life of a tree, it seems probable that records on individual trees will always be necessary.

Size of the Horticultural Plot

With the small fruits that are grown here, *e.g.* strawberries, currants and raspberries, as long as clonal material is available, individual plant records are seldom taken; since not only is the life of the plant shorter, but the difficulties of recording each plant are such that it can only be done for a comparatively small number. The value of uniformity trials, however, in which individual bushes or plants are recorded, cannot be over-estimated when starting trials on a new crop.

With strawberries, 100 plants has been found to be a convenient sized plot; with black-currants four bush plots are about the minimum, and generally speaking eight or ten bushes would seem to be about the best size for variety trials and up to twenty-five bushes for manurial trials.

The individual tree is always regarded as a plot for all tree fruits. This may of course entail difficulties in planning and management, especially since the adoption of modern methods of lay-out, which involve random distribution. In some cases modifications in design have had to be made to enable such operations as spraying, picking, etc., to be carried out more easily, and in others, *e.g.* manurial trials, it has been found necessary to make the plots considerably larger. In this case, however, whether in each differently manured plot the trees are of several different varieties or all the same, individual tree records must still be taken.

The pomologist is thus faced with the necessity of recording minutely the history of a large number of trees for a long period.

The weighing of the crop is undoubtedly the most important record to be considered, but the research worker cannot understand the cropping performance of his trees unless he has a sound knowledge of their vigour, since, even though the most vigorous tree does not necessarily produce the most fruit, the two characteristics are always closely connected. Again, just as the effect of a fertiliser may be shown in the straw of a cereal crop and not in the grain, so the effect of a system of manuring may be shown in the vigour of a fruit tree or bush, although no effect on crop is apparent.

The problem therefore arises as to how to record the vigour of a tree from year to year.

Measurements of Vigour

Few attempts have been made to record the vigour of small fruits until recently, when such measures as the cane length and number of canes to a stool of raspberries, or the spread of strawberry plants have proved exceedingly valuable.

The recording of the life of a tree begins in the nursery, the first notes to be taken being records of the size of each rootstock. Calliper measurements of diameter are a useful measure of size and sometimes notes are taken on the amount of root growth.

After the tree is budded and has made its first year's growth, it is usually planted in its permanent quarters and at this time records of height, weight and girth of stem are taken.

The girth is usually taken at a fixed height and this marked on the stem with white paint so that it can be measured in exactly the same place in after years.

After planting, several different vigour records may be taken. These include: measurements of annual twig growth, girth of stem, average height, average spread, and in some cases leaf size.

There has been considerable discussion as to which of these records best represents the vigour of the tree. After the examination of the records of very large numbers of trees at East Malling, the writer has come to the somewhat reluctant conclusion that no single record can be considered the best; since, though nearly all these measures of vigour are positively correlated where the trees are treated alike, differences in rootstock, manuring, pruning, etc., are all liable to upset the relationship between the different tree characters.

The cropping of the tree also has an effect upon its vigour. Increase in cross-section is generally low in a year of heavy crop, but the twig growth may not be adversely affected. Again, a heavy crop may weigh down the branches, so as to give a large increase in spread but an actual decrease in height.

The research worker therefore has to consider very carefully the possible results of an experimental treatment before deciding which measures of vigour should be recorded.

Measurements of Productivity

The first measure of productivity of interest to the research worker is the amount of blossom produced by the tree. This record is rather laborious to take, but in many experiments it is essential that some information concerning time of opening, amount of blossom, and proportion of blossoms which set fruit should be obtained. In seasons when weather conditions cause the failure of the crop, blossom records are the only measure of productivity available.

At harvest time weight of fruit is the first obvious record. It is often necessary, however, to supplement this by a count of fruit. Size, colour, quality and cleanliness are other characters of the fruit which must generally be recorded at this time.

Other Records

The above list of records of growth and fruiting may be regarded as more or less of a routine nature. In addition it is necessary to take records of other happenings in the life of the tree as often as possible. The incidence of a particular disease or pest, suckering from the roots, and leaf scorch, are some of the occurrences which need to be recorded from time to time.

Finally there are the personal notes of the research worker or his assistants. The small happenings on the plantation, this or that accident to a particular tree, the throwing out of roots from the scion, local variations in fertility or depth of soil, spray injury, etc., should all be carefully noted, for therein lie, not only the inspiration for these routine records, but also the means of interpreting them in after years.

Organisation of Recording

The taking of routine records and management on large scale experiments, wherein the trees or plots are often scattered at random, need very careful organisation, especially when unskilled labour has to be used. The first essential is the careful demarkation of plots.

Where the individual trees are scattered at random, as has been noted before, modifications in design are often necessary to make management possible at all. For example, in a large scale pruning trial on pears recently laid down at East Malling, nine varieties of pear are to be pruned in four different ways. The complete randomisation of these trees would make operations at pruning time very difficult and thus the pruning treatments were arranged more or less systematically, there being two treatments alternately in each row. Apart from this reservation, the thirty-six combinations of variety and treatment were arranged in randomised blocks in the ordinary way.

Although these varieties will be fairly distinctive by the time they are cropping, they will none the less have to be very carefully labelled to avoid errors.

The use of different colours in plot labelling, etc., may be very valuable, especially where illiterate labour has to be used. For example in a manurial trial involving the three constituents nitrogen, potash and phosphates, all the trees which get nitrogen could be given a red band, potash, a blue band, and phosphates white. It is possible to buy different coloured paper bags into which the different fertilisers can be weighed out, and thus the difficulty of application of manures on randomised plots considerably eased. One worker abroad marked his plots in this way and gave his native pickers armlets corresponding to the plot in which they were supposed to be picking, thus making the task of supervision much easier.

The necessity for guard rows in manurial trials, where larger plots are used, may also be turned to advantage in this connection, for a distinctive variety, or even another species of tree altogether, materially helps to mark plot boundaries in large orchards.

The use of Labour Saving Devices

Some of the records of vigour and productivity which have been alluded to are, to say the least, difficult to obtain ; and often the use of some labour-saving device may make them possible, where otherwise they would have to be abandoned.

Annual twig growth is certainly one of the most valuable vigour records, but also the most difficult to record. On older trees, indeed, it can only be taken for a few trees, since it may take up to an hour to measure a single tree. In most cases, therefore, other records, which give an idea of this character, such as weight of prunings, number of shoots, etc., have to be taken instead. Where growth is actually measured, a considerable saving in time and labour is made by the use of a field telephone connected to the laboratory, where the measurements can be recorded and added up on an adding machine in one operation.

A very good idea of the number of shoots on spur-pruned trees can be obtained from the number of cuts made in pruning. This can be obtained very easily by the use of a pair of secateurs with a counter attached.

At harvest time much labour is saved and many errors avoided, if as many records as possible can be taken in the field immediately the fruit is picked and before it is removed from under the trees.

The number of fruits can be counted automatically by means of a counter attached to a picking bag, so that each fruit is recorded as it enters the bag. The fruit is then weighed upon a portable spring balance adjusted for the weight of the bag.

Grading for size and colour must, of course, be done in the packing shed, and here an ordinary commercial grader can be used.

Sampling

The taking of samples in horticulture probably presents greater difficulties than with agricultural crops, since it is very difficult to take a definite sample from a tree. Thus a single branch will vary tremendously in size ; and a definite length of branch measured off and recorded will be a different proportion of the tree according to the size of the whole.

For this reason sampling with the object of measuring the actual crop or size of a tree has not as yet progressed very far.

On the other hand, qualitative characters have been successfully sampled in recent years. Thus by sampling the prunings from spur-pruned trees the weight of wood per unit length has been obtained and this in combination with the actual weight of prunings from individual trees and the measurement of a small number, has given a good idea of the annual wood growth.

Counts of blossoms and fruits upon measured samples taken at random have been used successfully to determine the percentage set of fruit; and daily counts of flowers open on similar samples have been used to determine the time of blossoming.

Again, random samples have been used to determine size, colour and such disease records as apple scab.

It is most important to remember, however, that such samples must be taken directly from the tree; it is very difficult to pick out a random sample from a box of fruit after it has been picked.

The Management of Fruit Experiments

There is one final point which must be included in any discussion of technique, and that is the proper procedure in such operations as pruning, spraying, manuring and thinning of fruit, where a large number of different kinds of tree are present in the same experiment. The problem arose recently, for example, in a manurial trial wherein nine different kinds of apple tree are included.

Should all these trees be pruned as the grower would prune them, to the best advantage, *e.g.* the large trees tipped more heavily than the small, or should the same proportion of wood be removed from all trees?

At first sight it would seem that the former course would be the better, since any single method adopted would probably be best for none; but if that were done what would happen when all the different effects of the manuring came into play? The larger kinds of tree when unmanured might actually require the same pruning as the smaller kinds when fully manured.

It seems therefore that it is better to treat all kinds of tree as far as possible alike, whether pruning, spraying or thinning; and this is the course generally adopted.

It is not possible to describe in detail all the methods of recording and difficulties with which the pomologist has to contend; but the experiences at East Malling have been described in a paper recently issued.[1] However, the best ways of overcoming the difficulties which

[1] "Field Experiments in Horticulture," by T. N. Hoblyn. Imperial Bureau of Fruit Production: Technical Communication No. 2 (1931).

arise are those which are evolved by the research worker himself and which suit his own particular conditions.

PRACTICAL DETAILS OF EXPERIMENTATION ON ORDINARY COMMERCIAL FARMS

By H. V. GARNER

Rothamsted Experimental Station

Experimental plots on the commercial farm call for labour of an unusual kind at busy seasons of the year, and there is sometimes a certain amount of difficulty in inducing farmers to co-operate with experimental stations in field work. Our experience has been that having once accepted an experiment the disturbance of farm routine has proved less than might have been expected, and farmers have been willing to retain experiments for a period of years. This co-operation is beneficial to both parties; for the research institute can test its findings under different soil and climatic conditions, while the farmer has the advantage of the only fertiliser experiments which he can interpret with confidence, namely those carried out on his own soil. Accordingly the aim should be to design experiments which have sufficient practical bearing to appeal to the farmer, while also providing information on more general questions.

Hitherto most experiments on commercial farms have been on single or in some cases on duplicate plots, but more elaborate arrangements involving higher replication are required to bring out small differences. Moreover, the finer points, if definitely established, are well worth the farmer's notice. The purpose of this paper is to describe the methods used in carrying out modern replicated experiments under the ordinary conditions of commerical farming.

The work has been done during the last few years by members of the Rothamsted Staff. The experimental centres cover a wide radius up to 150 miles from headquarters. Transport has been by train and hired car and the equipment no more than can be carried as passengers' luggage. Usually two supervisors have been required.

Co-operation with the Farm Staff.—The essential for the success of experiments of the kind under consideration is the close co-operation between the experimenters and the farmer. The nature of the trial should be fully discussed with him, and any modification which he may suggest should be adopted if found to help the main object in view. The interest and help of the farm foreman is equally necessary, for the arrangements with regard to labour and equipment are usually in his hands, and much time can be saved on the farm if this part of the programme is well looked after. Also it will usually rest with him to see that the experimental area is properly drilled and kept free from any manures which are to be applied to the rest of the field.

Choice of Land.—The experimental error is largely determined by soil variation, and great care is necessary in choosing the site for the experiment. Here again the help of the farmer and his foreman is essential, for cross-cropped or cross-manured areas, wet spots, and even the sites of old dung hills are not always apparent to outsiders. Minor soil variations which exist in every field can also be brought to the notice of the experimenters. A simple test for lime over the proposed area is a safeguard where sugar beet or clover is to be grown, and when an experiment is to remain for a series of years a soil analysis is used as a basis of the choice of a suitable site.

Plot Arrangement.—Latin squares involving four, five or six treatments are the forms preferred for soils of a rather variable character. These are sometimes modified by splitting the plots to take in another comparison, thereby doubling the number of plots without increasing the area, or by interweaving the rows or columns with two varieties of the same crop. If a larger number of treatments than six are to be investigated as in " balance " experiments, randomised blocks are used with four replications. The plots may be split in this form also.

Size of Plots.—Usually the main plots are $^1/_{30}$ to $^1/_{50}$ acre for both cereals and root crops. In root crops these are, if necessary, split into halves, for cereal into quarters (for harvesting by the sampling method). Labour considerations make the total area of more account than the number of plots, and for easy handling it should not exceed $^3/_4$ acre; indeed $^1/_2$ an acre is large enough if a heavy crop of roots and tops is likely to be grown. Up to the present, hay plots have been larger, $^1/_{10}$ to $^1/_{15}$ of an acre is the usual size.

Mixing Manures.—The farmer should be consulted about the scale of dressings, for his previous treatment of the land, *e.g.* the application of dung, will affect this question. Fertilisers of known analysis are always used and dressings worked out in terms of N, P_2O_5 and K_2O. When manures can be mixed at home and taken straight to the field this is desirable, though railway officials and taxi drivers seldom regard 5 cwt. of manure as passengers' luggage. The alternative is to send the manures in bulk to the farm well ahead of sowing time and weigh, mix, bag and label on the farm. A portable spring balance reading 20 lb. by ounces is quite satisfactory for this work. The caking of manures on storage renders the despatch of ready mixed manures by goods train rather risky, but it is hoped to overcome this difficulty in future.

Laying-out of Plots.—The experimental equipment consists of chain, crosshead, sighting posts, pegs and string. Plots are set out square, with their length parallel to the line of drilling. Pegs are put in all round the outside boundary and the whole strung along and across to define the individual plots for manure sowing.

Sowing Manures.—After checking the bags on the plots—a very necessary precaution—the sowing is done by hand. Farm hands do this job quite well if the nature of the work is explained. The stringing renders the chance of mistakes very small, and the supervisors can usually lend a hand with the work.

Locating the Plots.—Two of the corners of the experimental block are permanently fixed by driving in at each of them a stout peg level with the ground. The distance from these sunk pegs to substantial posts in the field boundary is ascertained. As a safeguard against the loss of these posts, measurements are taken which enable them to be picked up from permanent land-marks (gateposts, trees, etc.). In rootcrops counting the rows from a definite point is an aid to picking up the experimental area. On leaving the experiment it is necessary to define the corners with conspicuous posts in order to avoid the area being run over by the farmer's own manures. The subsequent operations on the experimental area are now carried out exactly as the rest of the field and its presence is no hindrance to the farm.

Field Observations.—A detailed plan of the experiment is sent to the farmer and his notes and comments on the action of the manures are of great value. Special observations may require a visit from headquarters.

Harvest.—More help is required from the farm at this stage than at any other. A few days' notice of the approach of harvest is necessary in order to enable the plots to be separated from the rest of the field and also separated from each other. Swedes and sugar beet are separated by strings, and the farm staff allowed to pull and top them on the plots. Here again little supervision is necessary when the job is understood. Potato plots are separated by digging 3-foot lanes across the rows and one picker for each plot along the frontage is made responsible for the potatoes spun or ploughed out in his area (defined by stakes). If potatoes are hand dug only stringing across the rows is necessary. Potatoes are sacked up and weighed on a corn scale. Beet and swedes are weighed in a pignet on a tripod roughly constructed on the farm, using a 200 lb. spring balance.

Cereals are harvested by the sampling method, which is the subject of a paper by Mr. Watson. Hay is cut by the farm methods when $^1/_{10}$ acre plots are used and separated on the ground by strings. It is made and cocked by hand on the plots and weighed as described for sugar beet.

Sampling.—Soil samples are taken from the unmanured plots in case analytical work seems desirable from the yield data. Sugar beet is sampled for sugar, and hay for dry matter, every plot being treated independently. This is quite an appreciable addition to the labour of harvesting.

Labour Considerations.—The amount of work entailed by replicated experiments is regarded as a serious matter not only by farmers but also by many experimenters. The operations on the experiments carried out in the last few years have been timed in order to obtain a rough idea of the labour involved. The data are only to be taken as a very rough guide, for weather and crop conditions give rise to big variations. Since an experiment is seldom less than 32 plots, times have been worked out on this basis. The figures refer to time spent actually on the operations. An allowance must be made for the necessary preparations for starting and for clearing up.

AVERAGE TIMES REQUIRED FOR AN EXPERIMENT OF 32 PLOTS

Mix manures .	2 men	1½ hours	Plots $^1/_{50}$—$^1/_{100}$ acre
Mark out plots .	2 ,,	1¼ ,,	,, ,, ,, ,,
Apply Manure .	2 ,,	2 ,,	,, ,, ,, ,,
Locate Plots .	2 ,,	½ ,,	,, ,, ,, ,,
Hay Harvest—			
Weigh out of cock, and sample .	4 ,,	9 ,,	,, $^1/_{15}$ acre
Cereal Harvest—			
Sample barley .	2 ,,	6 ,,	,, $^1/_{200}$,,
Potato Harvest—			
Sack on plots; plough and 8 pickers .		7 ,,	,, $^1/_{50}$—$^1/_{70}$ acre.
Sack on plots; 6 forks and 9 pickers . .		8 ,,	,, ,, ,, ,,
Weigh sacks on plots . .	4 ,,	2½ ,,	,, ,, ,, ,,
Sugar Beet Harvest—			
Weigh roots and tops and sample on plots . .	3 ,,	10 ,,	,, ,, ,, ,,

Standard Error obtained.—The feature of replicated experiments is the reduction of experimental error and its valid estimation. It may be of interest to summarise the magnitude of the significant differences between treatments as found in recent experiments on the above lines. By the usual convention a significant difference is one exceeding three times the standard error. The figures are as follows :—

MAGNITUDE OF A SIGNIFICANT DIFFERENCE (3 TIMES THE STANDARD ERROR) BETWEEN TREATMENTS IN RECENT REPLICATED EXPERIMENT ON ORDINARY FARMS.

	Number of Experiments	*Significance Difference. Cwt. per Acre*	*Per cent. of Mean Yield*	*Approximate cash value of the Significant Difference*
Hay [1]	11	3·8	13·7	6s. 6d.
Barley { Grain	3	2·9	18·8	20s.
Barley { Straw	3	2·5	15·7	3s.
Swedes	2	23	8·0	17s. 6d.
Sugar beet { roots	5	20	12·1	40s.
Sugar beet { tops	2	21	13·2	8s.
Potatoes	10	20	8·3	80s.

[1] By sampling methods.

It is sometimes urged against replicated experiments that they are needlessly complicated for ordinary purposes, though they may reveal points of academic interest. The cash value of the above detectable difference are in most cases quite considerable and represent returns which would cover any reasonable expenditure on manures.

The technique is certainly not too precise for ordinary purposes, and it is doubtful whether a cruder one could give results that could not be obtained equally well by observation.

Scope of the work.—With the limited staff available the amount of work which can be undertaken is small, the chief difficulty being its seasonal character and the fact that the farms are widely scattered. To obtain information over a wider area suitable schemes having a common basis and involving the co-operation of workers at other centres are required. A beginning has been made in this direction, but much remains to be done before the soil types of England will have been carefully examined in their responses to manurial treatment.

https://doi.org/10.23637/ERADOC-1-214

METHODS OF ESTIMATION OF CROP GROWTH AND YIELD

By D. J. WATSON

Rothamsted Experimental Station

Field experiments are usually laid down as yield trials, in which often the only recorded results are the final weights of produce obtained from the differently treated plots. Such results are of course of great value in determining agricultural practice, but their scientific value is much increased if it is possible to interpret the differential effects of treatment in terms of physiology. The analysis of yield in terms of physiological processes in the plant presents many difficulties, because of the large number and the complexity of the variable factors inside the plant and in the soil and climate which are involved.

Some information can be obtained by careful visual observation of the crop throughout the growing period. In this way differences in the amount of growth, for example, or the incidence of disease may be detected, but it is always difficult to make numerical estimates of such observed differences, and any estimates made are apt to be unreliable, depending as they do on the experience and preconceptions of the observer. It is obvious that if an attempt at an analysis of yield is to be made, direct measurements throughout the growth of the crop are essential. To begin with, we may examine the simpler aspects of growth, such as growth in height, growth of leaf area or tillering. This is a comparatively simple matter, but the problem becomes much more difficult if we attempt to push the analysis a stage further, by investigating the fundamental processes in the plant of which growth is the expression. At the present time there is a scarcity of methods of measurement of these physiological processes, which are at the same time reasonably accurate and sufficiently simple and quick to be employed in field work, and which do not involve destruction or damage to part of the crop.

Whatever aspects of the growth of a crop in a field-experiment we may wish to study a further problem presents itself. Because of the very large number of plants on an experimental plot, it is only possible to make measurements on a small fraction of the whole, and a sample must be selected from the plot for this purpose. Since there is usually a high variability between the plants, the validity of conclusions derived from measurements made on a sample, depend on whether the sample is really representative of the plot from which

https://doi.org/10.23637/ERADOC-1-214

it is taken or not. A reliable sampling method is therefore essential for field study of growth and physiology.

Recently at Rothamsted a new technique of sampling has been evolved by Mr. Clapham,[1] in which use is made of the statistical method of the Analysis of Variance. This sampling technique was first used for making growth measurements, but it has been shown that it could be adapted for estimating the yield of a plot with sufficient accuracy. Its use for this purpose has been most fully worked out for cereal crops, but it also has been applied, perhaps less successfully, to root crops. It will be simplest to discuss first the principles involved in the technique as applied to the estimation of yield of cereals.

A good sampling method must satisfy two requirements. First, the sample must be sufficiently large and so distributed in the plot as to be representative of the whole plot. Secondly, and more important, the sampling must be carried out in such a way as to make a statistically valid estimate of the error introduced in taking the sample possible. Unless such an estimate of error can be made the validity of the conclusions drawn from the observations on the sample remains in doubt, and also it is not possible to investigate the causes of the sampling error, and then by suitably rearranging the method of sampling to reduce the sampling error. A statistically valid estimate or error can be made only if the sample from a plot consists of a number of parts, or sampling units, which are distributed at random in the plot, that is to say, if the sampling units are a random selection from a large population of such units. Older methods of sampling, as for example, those used by Engledow,[2] and by American workers [3, 4] have involved a systematic distribution of the parts of the sample over the plot, and though a systematic arrangement reduces the labour involved in sampling, it is open to the objection that no valid estimate of error can be obtained.

Clapham took as his sampling-unit a metre-length of drill row. From a plot of about $^1/_{10}$th of an acre he cut thirty such metre-lengths, which were distributed at random. The yield of each of these metre-lengths was determined separately. The sums of squares of the deviations of the yields of each metre-length from the mean yield per metre-length, divided by the number of degrees of freedom, which is one less than the number of metre-lengths, and by the number of metre-lengths, gives the variance due to sampling, and the square root of this variance is the sampling error per plot. The sampling error worked out at about 6 per cent., and indicated that for an equal number of sampling units taken from a plot of $^1/_{40}$ acre, the usual size of the experimental plots at Rothamsted, the sampling error would be not more than 5 per cent. The standard error of plots of this size due to causes other than sampling is usually from 8 to 10 per cent., and the superposition of a further error of 5 per cent. would increase this only to from 9 to 11 per cent., so that the increased inaccuracy due to sampling was satisfactorily small. The yield of the whole plot was calculated from the mean yield of a metre-length, by multiplying it by the total length of drill row, which is easily calculated from the area of the plot and the distance between drill-rows.

Several refinements have been made in this original simple technique. The sample can be made more representative of the whole plot, by dividing the plot into a number of parts from each of which the same number of randomly distributed sampling units is taken. The procedure has the additional merit that the total variation between metre-lengths can be analysed into a portion representing differences between the mean yields of the parts of the plot and a portion representing differences between metre-lengths within the same part. The former portion of the variation may fairly be eliminated as due to differences in mean fertility between the parts of the plot, and consequently the sampling error calculated from the remaining variance is reduced.

The metre-length proved in some conditions to be too coarse a unit. This was shown by dividing the metre-lengths into halves, and harvesting each half separately. When this was done it was sometimes found that there was a significant correlation between the yields of the two halves of metre-lengths. That is to say, successive half-metres of drill-row were more alike than randomly distributed half-metres, and a given number of randomly placed metre-lengths was a less representative sample of the plot than twice the number of randomly placed half-metres. Doubling the number of randomly placed sampling units would of course increase the labour of sampling. A better solution of the difficulty was found by using a dissected metre-length of drill row. Various methods of dissection have been used. In one arrangement the metre was divided into quarters each of which was separated from its neighbour by half a metre of unsampled corn. In another the metre was divided into halves, which were taken from adjacent drill rows. With either of these arrangements the correlation between separate parts of the metre disappeared. A half-metre length of drill row is then a satisfactory unit for sampling a cereal crop, and it may be profitable in some cases to use a quarter-metre. Smaller units than this would be impracticable, as Professor Engledow [2] has pointed out, owing to the increased importance of end errors. American workers have used the rod-row ($5\frac{1}{2}$ yards) and the square yard as units, but these are much too large.

The general principles involved in Clapham's sampling technique may therefore be stated thus :—

The *sample* on which an estimate of yield of the plot is based consists of the aggregate of a number of *sampling units*, which are distributed at random over the plot. The *sampling unit* itself may be divided into a number of *units*, which are arranged systematically within the sampling unit. The particular systematic arrangement adopted is determined by two considerations. A good arrangement tends to make the sampling-units themselves representative of the plot, and therefore like one another, thereby reducing the sampling error. Also the arrangement must be such as to reduce as far as possible the labour of sampling.

The estimate of sampling error is derived from the variation between sampling units, and in order that an estimate of error may be made, at least two sampling units must be taken per plot. The size of the sampling error is influenced by the number and distribution of units within the sampling unit.

The accuracy of the estimate of sampling error depends on the number of degrees of freedom on which it is based, that is to say on the number of sampling units. For assessing the yield of a single plot it is therefore necessary to take a fairly large number of sampling units, depending on the size and uniformity of the plot, but probably not less than ten would need to be taken. When, however, a large number of plots, such as form a modern replicated experiment, is being sampled, each plot can be made to contribute to the estimate of sampling error, if it can be assumed that the variation from sample to sample is of the same order on different plots. There is considerable evidence to justify this for plots which are manured at ordinary agricultural rates, but it would not be true for a very wide range of manuring such as is often given in pot experiments. If the assumption can be made, the number of sampling units per plot may be cut down, often to as low as two per plot. There is then one degree of freedom per plot for estimation of sampling error, and if there are n plots, the sampling error for the plots taken together is based on n degrees of freedom. Since n would not usually be less than 16, and often considerably more, a sufficiently accurate estimate of error is obtained.

If the number of sampling units per plot is reduced in this way, it is necessary to increase the number of units in the sampling unit, to sample the plot effectively. For example, in a 4×4 Latin square experiment on barley, four sampling units were cut from each plot of $^1/_{40}$ acre, each consisting of ten half-metre lengths of drill arranged according to a simple systematic scheme. A considerable saving of labour is obtained by using a small number of large sampling units, since the number of random placings is reduced, and the number of threshings and weighings is also decreased. It is not necessary to determine the yield of each unit of the sampling unit separately, unless it is desired to test whether the systematic arrangement of units within the sampling unit is a satisfactory one.

The size of sample to be taken from a plot depends on the crop and on the accuracy required. It can only be determined by direct experiment. For cereals it has been found in a considerable number of trials that a sample of 20 to 30 metre-lengths per plot of about $^1/_{40}$ acre is sufficiently large, and gives a sampling error of between 5 and 6 per cent., which causes an increase of between 1 and 2 per cent. in the experimental error per plot, by which treatments are compared.

The technique has been applied to the estimation of yield of a potato crop.[5] A single plant was taken as the unit. Two sampling units were taken from each plot, each consisting of every twentieth plant in the rows of the plot, starting at a plant selected at random from amongst the first twenty plants. Thus each sampling unit was distributed over the whole plot and in itself representative of the whole plot. This is a good example of the way in which the advantages of a systematic arrangement in the saving of labour can be combined with the element of randomness which is necessary for a valid estimation of error.

The real test of the efficiency of a sampling method is made by comparing the yields estimated by sampling with the yields obtained by harvesting and weighing the whole produce of the plots. This has been done in a number of experiments,[6] and it was found that little information was lost by sampling, and the results obtained by the two methods were substantially the same.

We may turn to the use of the sampling technique for making observations and measurements during the growth of a crop. The procedure of selecting a sample for this purpose is precisely similar to that used in sampling for yield. But since counting and making measurements on the plants in a sampling unit takes very much longer than merely cutting out the unit and tying it up, as is done in sampling for yield, it is usually necessary to reduce the size of the sample considerably. Alternatively in a large replicated experiment, labour may be cut down by sampling from only a selected few of the replicates. This leads usually to an increased and less accurately determined standard error for treatment comparisons, since the number of degrees of freedom on which it is based is reduced. It is probably preferable therefore in most cases to reduce the number of sampling units taken from each plot rather than to reduce the number of plots sampled.

Reducing the size of the sample, of course, leads to an increased sampling error, and a loss of accuracy. But since the sampling error can always be calculated, and the significance of the results obtained from the samples determined, there is no danger of drawing unwarranted conclusions. In some recent counts of plant number and tiller number on the plots of a wheat experiment, where the crop was not abnormally uniform, 3 metre-lengths of drill row were counted in each plot of $^1/_{50}$ acre. The sampling error per plot was just under 12 per cent., but highly significant differences were found for the differently treated plots. There were 48 plots in the experiment, so that in all 144 metre-lengths were sampled, from an area of about one acre. Although this is small compared with the size of sample taken for estimation of yield, it represents a larger sample than has often been used in studies of crop development. Engledow,[2] for example, in his "Census of an acre of corn," took 100 foot-lengths of drill from an area of one acre, and concluded that this was a sufficiently representative sample.

The procedure of sampling in the field is very simple. A list of random placings for the sampling units is prepared beforehand. Usually each placing is fixed by two numbers, the first being the number of rows along one side of the plot, and the second the number of paces into the plot. After the numbers have been selected they are arranged in such an order that in observing successive sampling units the observer moves steadily over the plot, from one side to the other, and the amount of trampling is thus reduced. The sampling unit is measured out by means of a rod, in the case of a cereal crop, which is placed along the drill-row. If the sampling unit is a small one, a dissected metre-length, for example, it can be fixed by one placing of the measuring rod, but if a more complex sampling unit is used, a number of placings are necessary, following a systematic distribution from the randomly determined starting point. In sampling for yield, each sampling unit is cut out by means of large scissors or shears, and tied up and labelled with the number of the plot. The ears are protected by enclosing them in a paper-bag, which is perforated with holes, which are small enough to prevent the grains passing through them, in order that the ears may be adequately ventilated and the growth of moulds prevented. When all the sampling units have been taken from a plot, they are tied up together and brought back to the laboratory, where they are stored until it is convenient to weigh and thresh them.

The advantages which the sampling method gives may be summarised as follows :—

Many of the errors which are involved in large scale harvesting of cereals are avoided by sampling, as for example, losses of grain in the stook and in the stack. Inaccuracies due to weighing of weeds as straw are eliminated. Edge-rows can be discarded without the necessity of cutting them out. The bulked produce of the independently located sampling units form an excellent sample for chemical analysis. Smaller plots can be dealt with by the sampling method than by ordinary farm methods, and this is of great importance, since for a given experimental area, greater replication can be obtained by reducing the size of plot, and the accuracy of the experiment so increased. Where large scale machinery suitable for dealing with small experimental plots is not available, the sampling method may be used, and the problem of harvesting complex field experiments at farms some distance from the organising centre can be solved. Finally, since sampling in some form is necessary for the study of crop growth and development, a statistically sound method of sampling is indispensable.

REFERENCES

[1] Clapham, A. R., *J. Agric. Sci.* **19,** 1929.
[2] Engledow, F. L., *J. Agric. Sci.*, **16,** 1926.
[3] Arny, A. C., and Garber, R. J., *J. Amer. Soc. Agron.*, **11,** 1919.
[4] Arny, A. C., and Steinmetz, F. H., *J. Amer. Soc. Agron.*, **11,** 1919.
[5] Wishart, J., and Clapham, A. R., *J. Agric. Sci.*, **19,** 1929.
[6] Clapham, A. R., *J. Agric. Sci.*, **21,** 1931.

THE IMPORTANCE OF FIELD EXPERIMENTS

By T. H. J. CARROLL

Imperial Chemical Industries

During the last few years the number of artificial fertilisers placed on the market has considerably increased. In particular Imperial Chemical Industries has placed before the farmer a number of concentrated fertilisers such as have not previously been available in this country.

It is of the greatest importance to the fertiliser industry that it should know as accurately as possible the value of its products.

Numerous institutions and agricultural stations in this country are engaged in establishing the general value of fertilisers. There are some, however, who do not care to include the new concentrated fertilisers in their programme of work because they were made by one commercial firm.

Since, however, it is essential that the manufacturers must have as accurate an estimate as possible of the value of the fertilisers they offer to the farmer it is obvious that they should encourage such institutes to experiment with fertilisers, either by providing grants of money or in other ways. But this is not sufficient. The methods by which fertiliser experiments are carried out by such bodies must be carefully examined, and the results scrutinised so as to be able to judge to what extent these official experiments supply the fertiliser makers with reliable material for judging and for proving the importance of their products.

So far as one is able to judge from the published results of experiments by many agricultural stations the fertiliser industry is not being materially helped in its search for knowledge of the effects of its products. For instance published reports seldom state the number of replications used, the method of lay-out is rarely stated, and the degree of significance of the results is not stated. Moreover, information essential to the proper understanding of the trial, such as previous manurial treatment of the ground, previous cropping, etc., is usually omitted. More important still the economic aspect of the situation receives no attention as a rule, although a profitable return on money expended on fertilisers is almost invariably the sole reason for using them, and the real value of a fertiliser lies, not in increased yields so much as in the amount of profit which its application provides to the user.

Reliable investigations into the effects of fertilisers which may be expected under normal farming conditions are only possible by means of tests with fertilisers carried out on a large scale. Such tests should be carried out on a uniform plan, all the treatments should be replicated a sufficient number of times to allow of the results being statistically examined, and the trials should be so distributed that the effects of varying weather conditions can be studied.

The cultivation treatment of all the plots should of course be identical, and at each centre each operation should cover all the plots on the same day. The seed used at all the centres should, so far as possible, be the same, and the same methods of harvesting should be followed.

The object of such a multiple scheme of experiments should be to obtain an exact knowledge of the efficiency of the various fertilisers which contain one or more of the three principal plant foods. That is to say, one should be able to ascertain for the country as a whole the effects of nitrogen, phosphoric acid and potash which the farmer can count on getting in a normal season. At the same time it should be possible to find what combination of plant foods, and what quantities of them provide, the most *profitable* result.

https://doi.org/10.23637/ERADOC-1-214

The experiments should also be designed to ascertain the most favourable time of application.

Unfortunately we are still far from this ideal state of affairs.

Although thousands of so-called fertiliser experiments have been carried out in this country we have scarcely yet begun to collect reliable information in a scientific way. The farmer is not altogether to blame for not taking the results of previous fertiliser experiments into account when planning his fertiliser programme. He realises, better than we imagine, the limitations of the *ordinary* fertiliser trial.

If we are to arrive at certain and exact results we must be able to put a value on all those factors which influence the yield, including the one which is the object of the investigation, namely the fertiliser effect. The factor which it appears impossible at present to assess is the influence of the weather. The weather sometimes favours the plant and sometimes hinders it. It is not so much the weather over the whole season as the particular climatic conditions at critical stages of the plant's life. Weather conditions, therefore, being frequently the factor with the greatest influence on the harvest also determine the efficiency of the fertiliser in an important way. For instance, the fluctuations in the efficiency of fertilisers caused directly by weather conditions may be so considerable that, on the same site, the efficiency of 1 lb. of nitrogen during three successive years may vary between 20, 12 and 28 lb. of grain. It is evident that if the farmer intended profiting from the results of such a trial for one year only, he would arrive at false conclusions regarding the *average* fertilising efficiency of the material tested. As most of the published results of trials carried out in this country are due to demonstrations of one year only, the majority of farmers are right in rejecting these results as a basis of their future fertilising plans.

How then can we make the results of our fertiliser trials more exact so as to provide the farmer with a more reliable guide for his scheme of fertilising? We must first of all ascertain the *average* effect of the weather over a more or less extended period of years. This could be done by carrying out the experiment at the same place over several years, but such period would have to be at least five years and probably ten. But for propaganda purposes the fertiliser industry requires to be rapidly informed of the value of its products, and even five years is a long time to wait.

The alternative method is to have the largest possible number of trials in regions where the soil type is fairly uniform and cultivation methods similar. Weather conditions would generally vary in different districts of such regions, and the influence of the weather might be ascertained over one or two years by combining the results of such regional trials. The accuracy of such estimation will be all the greater if the number of trials is large and if the weather conditions vary to a considerable degree in the various districts of the region under consideration.

To sum up one may say that the routine methods of trials as at present largely practised do not meet our modern demands for accuracy. A search should be made for regions where the agricultural and climatic conditions are practically the same. In such regions a large number of experiments should be organised according to the same method. Only in this way can the real efficiency of fertilisers be ascertained. This is specially obvious when we are considering the effects of one particular plant food applied in different forms or at different times, where the effects may differ only to a very small degree and where the influence of the weather may be many times greater than such small differences due specifically to the fertiliser.

It is obvious that such a scheme of fertiliser experiments should be very carefully organised and carried out with meticulous care. This, however, is not impossible in practice, given a suitable staff of qualified workers. The number of agricultural institutes capable of carrying out such a scheme in this country is not sufficient to cover the whole country, and it will be necessary for fertiliser makers to engage in this work for themselves, but if such institutes as already exist could be persuaded to organise a large number of exact trials on a carefully conceived and uniform plan such work would go a long way towards providing the farmer with really reliable evidence of the practical value to him of modern chemical fertilisers.

DISCUSSION

Dr. E. S. Beaven (Warminster), whose well known half drill strip method has been used by the National Institute of Agricultural Botany in their variety trials for many years, opened the discussion. On the grounds of ease of manipulation in the field he strongly preferred a systematic arrangement to the randomised system used by the Rothamsted workers. The statistician was a good servant but a bad master. He considered that the justification for agricultural experiments lay in the degree to which they helped the farmer's pocket. Continental workers had been ahead of this country in carrying out replicated experiments with a uniform plan at numerous scattered centres. He instanced the early work done on this plan in Denmark with Barley varieties.

https://doi.org/10.23637/ERADOC-1-214

Dr. E. M. Crowther (Rothamsted) stressed the need for co-ordinated fertiliser expeiiments at a number of centres in order to provide material for chemical work on the available nutrients in soils. He mentioned the experimental rings as developed in Germany, where the financial support comes from the farmers themselves. An increasing amount of chemical work was now being done on the produce of experimental plots, and changes in the composition of the crop were frequently found even when no significant yield difference was demonstrated. For example, phosphates applied to hay had only a small effect on the yield in the first year, but in some areas almost doubled the amount of protein, and usually increased the content of P_2O_5. He mentioned recent lawn mowing trials where the degree of precision attained had been very satisfactory.

Mr. Holme (Kenya).—On considering the papers he was in some doubt as to the best way of conducting field work. He doubted if Imperial Departments could provide the personnel to carry out work on the lines described. The experimental work done in the past may not have been of a high enough standard for the laboratory worker, but it had produced substantial results for farmers who were looking for large increases and had no interest in improvements of the order of 2 or 3 per cent. He instanced the good service which the strip method of experimentation had performed in the Transvaal, and emphasised the importance of conducting a simple uniformity trial before accepting a new experimental area.

Dr. Wishart (Rothamsted) defended the complex experiments as conducted at Rothamsted on the grounds that it was the complex experiments which provided the smallest standard errors. The value of Dr. Beaven's half drill method lay in its high degree of replication (at least ten-fold). If the Rothamsted manurial experiments could be replicated to such a degree it would be an advantage, but at present, with the large number of treatments involved, this was not practicable.

Dr. Immer (U.S.A.) described the system of simple tests on ordinary farms carried out by Extension Agronomists with new varieties of cereals put out by the plant breeding stations. Single strips were sown at a large number of farms scattered over the State. Taken collectively, the information brought conviction and a 12 to 15 per cent. increase over the ordinary varieties was shown. The trials also demonstrated the parts of the State most suited to the new varieties.

https://doi.org/10.23637/ERADOC-1-214