THE STATISTICIAN AND PASTURE TRIALS

By A. C. GLENDAY, Applied Mathematics Laboratory, Department of Scientific and Industrial Research, Wellington.

In a paper read before the Royal Statistical Society in 1934 (1) appears the following quotation of 'an eminent Professor of Agriculture of an earlier date "Damn the duplicate plot; give me one plot and I know where I am." This certainly showed a realisation of the variability of plant material, but was also a confession of inability to overcome it and obtain valid experimental results. The present paper attempts to show to what extent the statistician has been able to provide adequate answers to the difficulties of dealing with the variable material in pasture trials. Much of what is presented here applies to other branches of science, and many of the principles involved are adapted from quite different branches.

The logical sequence of a pasture trial from planning to presentation will be adhered to in what follows

TYPE OF TRIAL

Under this heading only the following three types will be considered:

(a) Plot scale trial with mowing and return of clippings or fertiliser.

(b) Plot scale trial with common grazing of blocks and return of collected dung and urine in proportion to yields of individual plots.

(c) Paddock scale trial with grazing and movable cages for measurement.

(a) can only be used where the effect of the mower in changing the botanical composition and the loss of stock effects of trampling, selective grazing, and nutrient return are of minor importance. It is, however, of value in such situations as the comparison of species and strains in a new location or the comparison of levels of a fertiliser. (b) and (c) are the more realistic forms. Though (c) is the desirable one, the only real objection to (b) is the danger that the common grazing date imposed on a variety of treatments in a block may magnify or diminish the effects under trial. This is the main criterion that need be met before this type is adopted. In both types efforts must be made to move stock on from comparable feed and off at a similar state of appetite as at arrival to avoid transference of nutrients to or from the area, but this is more important statistically in the paddock trial. Also, as Donald (2) points out, the grazing of replicates rotationally ruins the trial statistically, for the replicates then become mere subdivisions of a single paddock. Other techniques which equalise the several plots of a particular treatment are open to the same objection.

In yield trials the statistician is not concerned with total annual yield of plots, but with indices of yield, so that constant errors in all yields through plot cutting heights differing slightly from grazing height or percentage errors in all yields through the omission from the record of growth during grazing are of no moment provided differences in these omissions do not reflect treatment effects; for it is easily shown mathematically that the changing of all yields by a constant or by a percentage, or by both, leaves the analysis and conclusions quite unaltered.

Thus far the decision on type of trial is made on practical grounds (space, facilities, object of trial) and not on statistical grounds. However, if the decision is still open at this stage, the statistical considerations in the next sections usually force a decision.

DESIGN OF TRIAL

The designs originally developed for field trials of arable crops are in common use in pasture work, though the tendency here is to use the simpler forms, randomised blocks and Latin squares, because of space, time and expense limitations rather than from any virtue of the simpler forms. In this connection support is given to Fisher (3), who wrote "No aphorism is more frequently repeated in connection with field trials than that **we 'must' ask Nature few questions,** or ideally, one question at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought'.out questionaire, indeed, if we ask her a single question, she will often refuse to answer until some other topic has. been discussed."

If. two factors each at several levels. are to be investigated in pasture work, the temptation is to do a simple trial using one level -of the second factor only with all levels of the first. The trial is repeated later with. a new. level of the second factor, but the same levels of the first factor, and so on. Relating the results of such trials done at different times and under different conditions is often impossible and the important interaction of the two factors is missed altogether. The only way to tackle the problem is to lay out the more complex factorial or split-plot experiment. This occupies more space at any time, but the total space-time units (i.e., cost) will be considerably less and the results valid and efficient instead of nebulous. As Crowther (4) points out, quite 'elaborate experiments are required to decide the best combination of several factors and "isolated experiments testing miscellaneous treatments with moderate replication are foredoomed to failure:"

In passing, it is worth noticing that the split-plot design allows compact experiments -of the plot type to be laid down within a paddock 'trial, and that more complex factorial designs can -often be used without replication in the ordinary sense.

NUMBER OF REPLICATIONS

The "power" of an 'experiment, that is, the order of difference which can be detected as real at the customary probability levels, is a definite mathematical function of the variability of the material, modified by design in use, and the number of replications. Thus if the statistician can obtain reliable figures for the variability of the data expected from. a trial, he can assist materially in ensuring efficient experimentation, where efficiency means the obtaining of maximum valid information for minimum effort. Given the.. detection power the experimenter would like his trial to attain, the statistician can forecast the number.of replications required, thereby ensuring that effort is not wasted in unnecessary. work or a project undertaken with insufficient replication for. its purpose. Similarly, given the amount of work or space the experimentalist can afford to-allot to a project, 'the statistician can forecast the detection power attainable. For example, using recent estimates of variability obtained at Grasslands Division from a trial of each of the three-types mentioned earlier, we get the following:

Type of Trial	Coeff. of variation (annual totals)	Number of treatments	No. of replications to detect as real diffs of 5 per cent of general mean	Space Required Acres
(a) 1/400 ac. plots	3%	6	3	1/20
(b) 1/300 " "	4%	6	5	1/9
(c) 1/20 " "	7%	6	16	4.8

The number of replications required in particular situations can be a serious drawback to scientific work, but this is the price that must be paid for sure knowledge. There is adequate reason here for doing all that is possible at plot scale before launching paddock scale trials, as advised in a recent report by the American Society of Agronomy (5).

SIZE, SHAPE AND ORIENTATION OF PLOTS AND BLOCKS

Just as the most efficient design for a given area and set of conditions can be determined by the statistician, so the optimum size and shape of plots and blocks can be found. These change with circumstances of course, but the following general rules are valid.

(a) Size: For a given area (disregarding considerations of effort), the smallest convenient size of plot is the most efficient, for this decreases block size, which itself adds to efficiency, and also increases replication.

For a given amount of work (disregarding considerations of space), a larger size of plot is better. No rule can be given here except that the size of plot should limit the block to 1/10 acre or less.

In paddock work, of course, the smallest manageable units are the obvious rule. This will be referred to again in a later section.

(b) Shape: The shape of plot is determined by block shape. On land where there is *no* greater variability in one direction than in others, the block should be as square as possible, and where there is a marked trend in one direction, the block should be oblong with its shorter dimension in the direction of the trend. These precautions ensure that the block effects removed in the analysis of the results account for as much of the variation as possible.

Plots should therefore be. long and narrow, with their lengths in the direction of the greatest fertility trend to ensure minimum bias in treatment estimates. Note that in the Latin square the plots should approach squareness, and that. there is little or no gain in efficiency in the Latin square if the overall shape of the layout is far from square.

In paddock work the smallest convenient size implies a square shape of plot, the plots in a block being arranged as square as possible.

(c) Orientation: Apart from considerations in (a) and (b) above the orientation of the blocks is completely at will. The relative positions of the blocks is also at will, and the blocks need not be contiguous provided the treatment differences do not change radically from block to block. Should two blocks be on different soil types, say, a treatment x soil type component of variance is incorporated with the residual variance determining the trial's accuracy, and the efficiency of the trial deteriorates markedly. Results from any trial can only be generalised to the type of country and conditions of which the experimental plots can be considered a representative sample. Restricting the whole layout to a small area increases precision, but only of conclusions applying to that soil type, etc. Broadening the area included broadens the generalisations possible. from the results,. but brings with it the necessity for more replication, as the unaccounted variation will inevitably increase. It may even -compel the adoption of a more complex design so that treatment x soil type and other interactions can be elucidated,

Overall block differences in yield do not matter, as these cause no bias in the estimates of treatment effects, and are eliminated entirely in the course of the mathematical analysis.

RANDOMISATION

Within the restrictions of the design the allotment of the treatments to the plots of the blocks must be at random. As Fisher (6) and Yates (7) showed, the positive correlation of yields from neighbouring plots destroys the theoretical basis for the common form. of analysis if -applied to systematic layouts, and the estimates of error and tests of significance are liable to be vitiated. Randomisation 'allows the analysis to proceed as though the difficulty did not exist. The loss of some ease of demonstration in a field trial is unimportant by comparison with the importance of validity of results.

SAMPLING

In none of the grazing trials considered here are animals used for yield determinations. Thus sampling must be used to secure all data.

The optimum shape of sampling unit for any 'purpose can be determined statistically, and in general this should be long and narrow for maximum efficiency. For a fixed area of sample a large number of small samples is better than a single large sample.

The major difficulty in sampling is number of samples to be taken. When a statistician is available good use can be made of a sequential sampling system. In this, two or more samples are first taken from each plot or paddock. From these, estimates of plot and sampling variances can be computed, and the increase of accuracy for any further number of samples or the number of samples required to reduce errors to stated limits can be forecast. By this means the most efficient use is made of the available labour. It is found that in cases where the within plot or sampling variation is high in relation to the between plot variation, the precision can be increased by further sampling, but when the sampling variation is low in relation to the plot variation, little or nothing can be gained by further sampling. For a stated sampling effort the greatest efficiency is attained when only one sample is taken from each plot, but the maximum number of replicate plots is sampled, Replication of samples within plots is never a substitute for replication of plots.

The sites for samples of any kind should be selected at random if valid tests of differences are to be obtained. It is admitted that skilled workers like Sears (8) may be able to get accurate estimates of a plot yield or composition by taking two "typical" samples, but this is doubted by others like Klapp (9). In any case, the objection of the statistician is embodied in the following question: Which is better, a figure of quite unknown bias or repeatability on which no claims or tests can be based, or one with absolutely no, bias and for which valid limits of error can be stated? In brief, the "typical" sample does not allow a test except on the assumption that every "typical" sample is an absolutely perfect record of its plot. The difficulty of the ordinary analysis of data from such samples, as pointed out by Dr Boyd in a paper by Linehan, Lowe and Stewart (10), is that only minimal standard errors are obtained. These give "significant" results too often where there may not be any real effects. That is, the statistical tests are biased.

A "systematic" sample is one in which sampling is done at equal intervals along a folded line covering the area in an irregular manner, as used by Nielen and Dirven (11) and advocated by Yates (12). This hardly violates at all the random requirement (that all possible samples in the, plot have an equal chance of appearing in the sample taken), and this is the sole criterion that need be used in judging whether the sample taken will allow of valid statistical testing.

The word "systematic" is often used wrongly to describe a "stratified" sample. In stratification the area is divided into equal sub-areas and one or more samples taken at random from each sub-area. This method gives greater accuracy where there are any trends across the plot, as these can be eliminated from the estimates of error, and is a necessity where the trend is pronounced. The gain in precision is analogous to that obtained by grouping plots into blocks in the randomised block layout, instead of having them scattered at random over the experimental area. This is the method recommended by Goodall (13) for point analysis determinations.

Sampling for botanical or chemical composition should be done from individual plots, not from the bulked plots of a treatment, if these measures are to be used to test differences in pasture components or chemical composition.

SPECIAL STATISTICAL TECHNIQUES

The only additional technique that can be dealt with in the space of this paper is that of covariance. The precision of any experiment is determined. solely by the size of the residual variance, that is, the variance of. the data unexplained by treatment, block, and other effects allowed for in the design. This often leaves much variation unexplained and an experiment of little "power." In many cases the remaining variation is further explainable by its relation to some other information which can be collected either before or during the trial. This additional information is called "concomitant" information, and allows a further reduction of error. Valid concomitant information is data whose correlation with final yields (after removal of assignable effects) is high, i.e., which represents some permanent feature of yields of uniformly treated plots on the area. Data affected by treatments must not be used.

In the ordinary field trial on arable crops running for one year it is usually inefficient to delay the experiment a year so that "uniformity" data can be collected with the object of increasing the precision of the main experiment to be run in the next year. It is usually better in these cases to trade time for space and to use more area for a shorter time:

In pasture trials the technique does give a notable gain in efficiency where the land used is very patchy, and in a recent pasture experiment on such land the detectable difference was reduced from 19.6 per cent of the general mean to 12.9 per cent by the use of data from two cuts before the start of the trial. This was so only because differences in the yields of plots on these two cuts represented permanent plot differences, but will not occur if the differences represent only temporary effects to be reversed later in the season. In long-term pasture trials it can often pay dividends to delay the treatments for one or 'even two growing seasons so that covariance can be 'used.

If the plots' to be used' have been differentiated -in fertility by past treatments, the technique **should** be used, supplying its own test of efficiency, as pointed out by Forester (14).

STANDARDS OF ACCURACY FOR COLLECTED AND PRESENTED -DATA

So that analyses and tests will have their full precision, it is necessary to retain 3-figure accuracy in the data collected, where coefficients of variation are in the usual range of 4 to 40 per. cent.' That is, if yield data for analysis average 25lb., weights should -be recorded to .1lb.

In final presentation, to make the conclusions as easy as possible to grasp without obscuring the tests or making the results appear more accurate than they really are,. means should. be rounded to 1/10 of their estimated standard errors, as stated- by Cochran and

Cox (-15). Thus, if a trial reporting yields averaging 15,000lb. d.m. per acre from 4 replications has a coefficient of variation of 6 per cent, these means could be rounded to the nearest 50lb.

CONCLUSION

In conclusion, in contrast to the opening quotation; I would like to leave with you a quotation from Yates in 1946 (16) "Statistical considerations now influence the planning of observational and experimental work, as well as the analysis of the results from such work. In consequence, statistics has begun to play a part in agriculture research which would have seemed incredible 25 years ago."

REFERENCES

1. Wishart, J. "Statistics in Agricultural Research" supp. J. Roy. Stat. Soc., 1, 1934.
2. Donald, C. M. "Pastures and Pasture Research" (U. of Sydney) 1946.
3. Fisher, R. A. "The arrangement of field experiments" J. Min. Ag. XXXIII, 1926.
4. Crowther, E. M. "A review of some developments in experimental agriculture" Proc. Britt. Soc. for Adv. of Sci. 1952.
5. American Soc. of Agronomy "Pasture and Range Research Techniques" Agronomy J., Vol. 44, 1, 1952.
6. Fisher, R. A. "Design of Experiments" (Oliver and Boyd 1948, Sect. 26.
7. Yates, F. "Comparative Advantages of Systematic and Randomised Arrangements" Biometrika Vol. XXX, 1938.
8. Sears, P. D. "The technique of pasture measurement" N.Z. J. of Sci. and Tech, A. Vol. 33, 1, 1951.
9. Klapp, E. "Methods of Studying Grassland Stands" Herb. Rev. 3, 1, 1935.
10. Linehan, P. A., Lowe, J., Stewart, R. H. "The Output of Pasture and its Measurement" Pt. II. J. Brit. Grass. Soc. 2, 3, 1947.

11. Nielen, J. F., Dirven, J. G. "The Accuracy of the 25 cm² specific frecjuency method" Rep. of Ag. Res., Wageningen, 56. 13;

12. Yates, F. "A Review of 'Recent Statistical Developments in Sampling Surveys" J. Roy. Stat. Soc. 109, 1946.

13. Goodall, D. W. "Some Considerations in the Use of. Point Quadrats 'for the Analysis of Vegetation" Aust. J. Sci. Res. B. 5, 1, 1952.

14. Forester, H. C. "Design of Agronomic Experiments for Plots Differentiated in Fertility by, Past Treatment" Iowa Res. Bull. 226, 1937.

15. Cochran, W. G. and Cox, G. M. "Experimental Designs" (Wiley) Sec. 3.33, 1950.

16. Yates, F. "The Place of Statistics in Agricultural Research" Ag. Frog. XXI, 1, 1946.

DISCUSSION

Q. (Lynch) : Mr Glenday stated that rotational grazing of grazing trials invalidates the statistical analysis of the results. In all our pasture mowing trials, whether rotationally grazed or not, we get an evening up of the trial as it progresses. Would Mr Glenday comment on this?

A. By rotational grazing the paddocks become progressively less variable. Finally if we begin with, say, 6 replications we reach the stage where we have in effect 1 plot instead of 6 replications. By using several dummy treatments it is possible to get significant results after several, cycles of this smoothing operation.

Q. (Lynch) : One is liable to fall into error in assuming that differences during the period of taking the uniformity data necessarily hold after the treatments are applied. Caution should be used when applying uniformity data to trial results.

A. The only data that can be used would be those representing permanent plot differences. On very variable land such as that referred to by Mr Lobb in the North Otago molybdenum trials pre-trial, cuts should certainly be obtained.