Report for 1937

Statistical Methods and Results

Rothamsted Research

being the more striking. On heavy clay loam at Rothamsted, although the qualitative composition is not seriously altered by mineral manures, some species are much encouraged and others are considerably reduced, the relative variation being influenced by season. The addition of nitrogen eliminates many species, and with heavy dressings a few grasses develop strongly at the expense of the rest. Nitrate of soda and ammonium sulphate do not encourage the same association of species on account of the difference in soil reaction. The acidity induced by heavy doses of ammonium sulphate much favours *Holcus lanatus*, but the addition of lime brings *Alopecurus pratensis* and *Arrhenatherum avenaceum* into predominance, leguminous and other plants being drastically reduced. With the neutral reaction induced by sodium nitrate *Alopecurus* and *Arrhenatherum* flourish without lime, shade being here more effective than liming in changing the proportion of species.

With organic fertilizers the yield may be reduced by heavy dressings of lime without very marked alterations in herbage composition. The response to lime is rapid, as the species affected usually show a variation in their relative proportion at the first cutting, although in some conditions the change may be delayed.

Certain species afford some indication of soil and manurial conditions. *Taraxacum vulgare* is prolific on well manured soils with a tendency to alkalinity; *Scabiosa arvensis* flourishes where potash is deficient and no nitrogen is applied; *Rumex acetosa* is possibly associated with scarcity of phosphate on soil which is otherwise well manured; while Leguminosae may form a third of the herbage where minerals without nitrogen are given.

STATISTICAL METHODS AND RESULTS
(Department of Statistics)

(a) DESIGN OF EXPERIMENTS

XII. F. YATES. "*The Gain in Efficiency Resulting from the Use of Balanced Designs.*" Supplement to the Journal of the Royal Statistical Society, 1938, Vol. V, pp. 70-74.

The comparative efficiency of a balanced design, which was actually used in a nutritional experiment on human beings, and other alternative and simpler designs, is assessed. It is shown that the balanced design is considerably more efficient than the others.

XIII. W. G. COCHRAN. "*Note on J. B. S. Haldane's paper 'The Exact Value of the Moments of the Distribution of χ^2.'*" Biometrica, 1938, Vol. XXIX, p. 407.

A discrepancy noted by Haldane between his and the writer's values for the mean and variance of χ^2 in a $2 \times$ n-fold contingency table with known expectations is shown to be entirely due to a difference in the definition of χ^2.

(b) ANALYSIS OF DATA

XIV. F. YATES and W. G. COCHRAN. "*The Analysis of Groups of Experiments.*" Journal of Agricultural Science, 1938, Vol. XXVIII, in the press.

When a set of experiments involving the same or similar treatments is carried out at a number of places, or in a number of years, the results usually require comprehensive examination and summary. In general, each set of results must be considered on its merits, and it is not possible to lay down rules of procedure that will be applicable in all cases, but there are certain preliminary steps in the analysis which can be dealt with in general terms. These are discussed in the present paper and illustrated by actual examples. It is pointed out that the ordinary analysis of variance procedure suitable for dealing with the results of a single experiment may require modification, owing to lack of equality in the errors of the different experiments, and owing to non-homogeneity of the components of the interaction of treatments with places and times.

XV. W. G. COCHRAN. "*Some Difficulties in the Statistical Analysis of Replicated Experiments.*" Empire Journal of Experimental Agriculture, 1938, Vol. VI, pp. 157-175.

The analysis of variance is now widely applied in interpreting the results of replicated experiments. Sometimes, however, a combined analysis on the original data has little meaning and gives misleading results, because the treatments have different variances. A numerical example is given to illustrate such a case.

These cases may be divided into two groups. (1) With yield-data, or whole-number counts of over 100 per plot, they occur very rarely, but may do so if some treatment differences are of the order of several hundred per cent., or if there is a partial failure of certain treatments or plots. The analysis is best carried out by omitting some treatments or plots. (2) With small whole numbers or percentages, the distributions tend to follow the Poisson and binomial types, respectively, and there is a known relation between the variance and the mean. Data of this type should be transformed before an analysis to a scale on which the variances are equal.

Three transformations have proved particularly useful in practice. (a) The square root, for whole numbers per plot between 10 and 100. If the majority of the plot-yields are under 10, one-half should be added to each plot-yield before taking the square root. (b) The inverse sine, i.e. the angle whose sine is the square root of the fraction, for percentages and fractions based on the ratio of small numbers. Percentages can, however, often be dealt with either by square roots, for small percentages, or by a direct analysis, for percentages from 30 to 70. (c) The logarithm, for distributions in which the standard error is proportional to the mean.

Numerical examples are worked, illustrating the use of each of these transformations and the way in which to present the results of the experiment.

A brief discussion is given of the analysis when the results consist of the number of plants in each of a number of grades (e.g., healthy, slightly diseased, severely diseased).

With factorial experiments in which the main effects produce large differences, the experimenter must consider what is the most natural definition of the independence of two factors, since the conventional test of interactions in either the original or the transformed scale may have little relation to this. A numerical example is given illustrating this point.

XVI.   W. G. COCHRAN. " *Recent Work on the Analysis of Variance.*" Journal of the Royal Statistical Society, 1938, Vol. CI, pp. **434-449.**

This review covers the period 1934-7. The principal topics summarised are experimental design, the discriminant function and its uses, the analysis of covariance and the use of transformations with non-normal data.

XVII.   H. FAIRFIELD SMITH. " *An Empirical Law Describing Heterogeneity in the Yields of Agricultural Crops.*" Journal of Agricultural Science, 1938, Vol. XXVIII, pp. 1-23.

The object of the paper was to investigate the relationship between the plot size and the variance of the plot yields. Using data from a blank experiment with wheat it was found that the regression of the logarithms of the variances for plots of different areas on the logarithms of their areas was approximately linear. A graphical review of variances, etc., reported in the literature for thirty-nine other blank experiments indicates that the results of most such experiments conform to the same law.

It is shown that the above law can be generalised (so as to be applicable to any size of field) by applying a certain adjustment to the regression coefficient b', so as to give a modified coefficient b applicable to an " infinite" field.

From this generalized relationship there has been deduced an expression to indicate average relative efficiencies to be expected for randomized block experiments with varying numbers of plots per block in a field for which the coefficient b is known.

A formula, which may be used to estimate the most efficient size of plot for any given experiment, has also been deduced. The cost of using plots of other than the most efficient size is indicated graphically.

(c) SAMPLING

XVIII.   W. G. COCHRAN. " *Crop Estimation and its Relation to Agricultural Meteorology.*" Supplement to the Journal of the Royal Statistical Society, 1938, Vol. V, pp. 1-25.

95

This is one of three papers read to the Industrial and Agricultural Section of the Royal Statistical Society, with discussion. The first paper, by Dr. Irwin, describes critically the official methods of crop estimation in the United Kingdom, the United States and India and outlines the early work of the crop-weather scheme of the Agricultural Meteorological Committee. In the second paper, by Mr. Cochran, the use in crop forecasting of the correlation of yields with weather factors and with previous measurements on the crop is discussed. A prediction formula is presented for wheat, based on the results of the wheat sampling observations under the crop weather scheme. This formula is as yet unsatisfactory in forecasting variations in yield from year to year, but may be improved by the inclusion of weather effects when sufficient data become available. In general, however, much research is still needed on the possibility of crop forecasting by this method.

The estimation of crop yields at harvest by taking small samples from a number of fields is considered and the results obtained from an extension to commercial crops of the wheat sampling scheme are presented. The chief difficulties here appear to lie in the selection of fields to be sampled and in a positive bias which persists in the sampling yields as compared with the harvested yields of whole fields. The use of a fixed panel of forms for each crop is recommended as the most practicable method of organising the work.

The third paper, by Dr. Wishart, sums up several points in the first two papers and describes the system of crop estimation in China.

THE SOIL

(Departments of Chemistry, Fermentation, and Physics)

(a) CULTIVATION

XIX. E. W. RUSSELL and B. A. KEEN. *" Studies in Soil Cultivation. VII. The Effect of Cultivation on Crop Yield."* Journal of Agricultural Science, 1938, Vol. XXVIII, pp. 212-233.

The yields of wheat, barley and mangolds were not appreciably affected whether the seed beds were prepared by ploughing and harrowing, by using the grubber (or cultivator) and harrowing, or by using the Rototiller, provided that the grubber and Rototiller were used for one year only. If used for several years in succession deterioration of yield sometimes sets in, possibly due to the increased weediness of the non-ploughed plots.

There was no advantage in ploughing deeper than 4 in. but it is advantageous to use the grubber or Rototiller deeper.

For spring-sown crops, cross-ploughing, subsoiling, or heavy rolling of the seed bed were without effect on the yield.

Spring rolling and harrowing improved the yield of winter wheat but had little effect on the yield of straw. Rolling alone produced a slightly increased yield of grain. The straw yield was increased by rolling but depressed by harrowing.

There was strong evidence that intensive hoeing of sugar-beet or kale is detrimental to the yield. Two to three hoeings appear to be ample.

XX. E. W. RUSSELL and N. P. MEHTA. *" Studies in Soil Cultivation. VIII. The Influence of the Seed Bed on Crop Growth."* Journal of Agricultural Science, 1938, Vol. XXVIII, pp. 272-298.

Crops germinate faster on the looser seed bed prepared by a Rototiller than on the more compact ones prepared by a plough or a grubber. The total number of plants that germinate is, however, the same for all treatments unless the land is too foul with weeds, when higher germination is obtained on the cleaner plots.

Cereals tend to ripen a little sooner on land that has been ploughed than on land that has been either rototilled or grubbed.

The roots of mangolds were longest and thinnest on the deep-ploughed plots and were always squatter on the shallow-tilled than on the deep-tilled plots. The roots were heaviest on the deep-ploughed plots and lightest on the rototilled plots. On the rototilled and the grubbed plots the depth of tillage had no effect. The plants on the shallow-grubbed plots seemed, however, to have no reserve of strength, for they could not make better growth if given more room, while those on the deep-grubbed plots could make