https://doi.org/10.29321/MAJ.10.A04117

Madras agric. J., **45 : (6) 223—230, 1958**

Replications in Field Experiments

by

T. A. GOVINDA IYER,
Agricultural College and Research Institute, Coimbatore.

Introduction : Research workers who have to deal with field trials are often faced with the problem of deciding upon the number of replications that should be adopted for a particular set of treatments. In most cases the number of replications is decided arbitrarily and this leads to a result where even large differences between treatment means fail to attain statistical significane. Any method by which the workers will be enabled to determine the number of replications necessary to render a predetermined difference, of say 10% between the treatments, to be statistically significant, would be very useful. An attempt is made in this paper to evolve such a method, utilising the data that were available from the Paddy Breeding Station, Coimbatore.

As is well known, the test for a significant difference between two treatment means is based on the value of 't', where

$$t = \frac{\text{D (difference between the means)}}{\text{standard error of this difference.}}$$

If 'n' is the number of replications of each treatment and "δ 2" is the expected error mean square as shown by previous experiments, then

$$t = \frac{D}{(\delta / \sqrt{n})\sqrt{2}} = \frac{D}{\delta} \times \frac{\sqrt{n}}{\sqrt{2}}$$

$$\text{Therefore, } n = \left\{ \frac{t \times \sqrt{2} \times \delta}{D} \right\}^2$$

By substituting the approximate values in this equation, D and δ usually as percentages of the general mean, it is possible to calculate the number of replications of each treatment necessary to prove significant any difference greater than D (Patterson 1939).

The value of standard error in an experiment expressed as a percentage of the general mean is called coefficient of variation (C. V.) i. e. the square root of the error mean square expressed as a percentage of the General mean. It is a measure of the experimental error involved in the experiment. A mean of such values from past experiments is what has been termed 'δ' in the above equation. The method suggested in this paper is therefore to obtain the mean C. V. from past experiments and to test its reliability for the use on hand.

Materials and methods: Yield trials on paddy conducted during the three years 1954 to 1956 at the Paddy Breeding Station, Coimbatore have been used for this study. Forty six trials had been conducted during this period. The plot size, number of paddy varieties, the general mean and the C. V. calculated in each case are tabulated in Appendix I.

The plot size has varied from about 30 to 100 sq. ft. the number of varieties from 3 to 10, and the general mean (expressed as yield per acre) from about 1000 to 4000 lbs. Thus, the normal variations in plot size, variety number and yield level may be considered to be represented in the three years' data dealt with in this paper. The reliability of the mean C. V. has been tested by correlation and regression methods as recommended by Goulden (1952).

Results and Discussion:

I. The arithmetic mean and standard deviation of C. V. and of the other characters are tabulated in Appendix. I.

The C. V. has a mean value of 12·57% and a standard error of 0·85% (standard error is standard diviation divided by the square root of 46). This value of 0·85% is a measure of the variability among *means* of C. V. that would be obtained on repeated sampling with a sample size of 46 experiments at each time. Assuming that such means will be normally distributed, which assumption is usually reasonable, as the sample size is fairly large, one can say that the limits $12{\cdot}57 \pm 2 \times 0{\cdot}85$ would contain true value on an average in 95 out of 100 cases. This also means that a mean C. V. estimated from a set of 46 experiments in the Paddy Breeding Station, Coimbatore has a probability 0·95 to fall within 1·7% on either side of the unknown true C. V. value. All these indicate the fair reliability of the culculated mean C. V. of 12·57% as an estimate of the true unknown value in the Paddy Breeding Station, Coimbatore.

Making use of this mean C. V. value, the number of replications required for an yield trial of strains at the Paddy Breeding Station, Coimbatore may now be calculated from the formula already given.

Thus, $$n = \left\{\frac{t \times \sqrt{2} \times C.\ V.}{D}\right\}^2 \qquad \text{—(A).}$$

where t — the reading from the 't' table for any desired probability (usually 0.05) and the number of degrees of freedom from which C. V. has been estimated (45 in this instance)

and d — the difference between the varietal means expressed as a percentage of the general mean. It is desired to detect statistical significance in the trial all differences exceeding this value. Let us assume it as 10% of the general mean. This would mean in other words, that the trial is desired to show statistical significance whenever the varietal ifferences exceed this limit of 10% of the general mean.

By substituting the values in the equation as below we get,

$$n = \left\{\frac{1{\cdot}96 \times 1{\cdot}414 \times 12{\cdot}57}{10}\right\}^2 = 12{\cdot}11.$$

According to the above, 12 would be the number of replications necessary for each strain in future yield trials at the Paddy Breeding Station, Coimbatore, if any difference greater than 10% of the general mean should be significant at the 5% level.

It will be noted from the equation that the lower the value of 'D' that is to say, when a higher level of precision is desired in the trials, a greater number of replications will have to be included. The level of precision, i. e. the value of 'D' must, therefore, be first decided upon and then the value of 'n' is obtained from the equation as illustrated above. If the number of replications estimated turns out to be too large to be adopted in practice, it may be reduced, with corresponding loss of information.

II. It might be noted from the above discussion that the mean C. V. value used should be a reliable estimate of the experimental error, because the number of replications estimated for a given level of precision depends mainly on this valne. The number of experiments averaged for its calculation and their representative character has already been mentioned. The fairly low standard error obtained for the mean C. V. is also a point counting in its favour. It remains to be seen, however, to what extent this mean C. V. has depended on some of the other important controllable field characters, such as plot size, number of varieties and general mean in the experiments considered, so that any adjustment necessary may be made to its value for reliability in its use. Workers elsewhere have found that these characters influence the mean C. V. to varying extent (Arnor & Nissen 1956).

In an attempt to study the nature and extent of dependence of the mean C. V. value on the different plot sizes, variety numbers and the general mean yields involved in the 46 experiments, the following coefficients of correlation (r) have been calculated for the relation between the four variables.

TABLE I.

The Coefficient of correlation (r) for the relationship between C. V. (y)., plot size (X 1), variety number (X 2) and yield level (X3).

	X1	X2	X3
Y	— 0·19	0·30 **	— 0·42 **
X1		— 0·20	0·10
X2			— 0·03

The study indicates that the C. V. has increased significantly with increasing number of varieties in the trial and also with decreasing general mean yield. There is also a trend for increase in C. V. with a decrease in plot size although the 'r' value obtained has not proved significant. A similar non-significant negative association has also been found to exist in the values for plot size and number of varieties. The other pairs of the 'Independent' characters have 'r' values not only non-significant but also very low. The 'r' values in all pairs including the significant values are low. Also none of the various pairs studied indicate any curvilinear trend.

It follows from these that the mean C. V. calculated from past experiments has depended on two of the field attributes, viz., the number of varieties under test and the general mean yield. The degree of dependence is however, small since only low values of 'r' have been obtained, and the data also do not suggest any curvilinear trends. The different plot sizes that were used in the trials and which may be considered to be the usual sizes in practice in such trials do not evidence any significant effect on C. V. meaning thereby that the plot sizes usually adopted at the station do not affect the reliability of the men C. V. obtained. However, the value of 'r' is not much lower than either the 't' value at 5% point or the other significant values calculated. Therefore the combined effect of all the three characters; plot size, number of varieties, and general mean, on C. V. has been estimated by calculating the Multiple correlation coefficient (R).

The multiple correlation coefficient (R) calculated is O. 51 and is highly significant. The corresponding coefficient of multiple determination value of 26% (R^2 expressed as a percentage) brings out clearly that only 26% of the total variation in C. V. values has been accounted for by the concomitant variation of all the three characters studied simultaneously. The combined influence of plot size, number of varieties and general mean on C. V. is thus only 26% of the total unknown influence to which it is subject.

A multiple regression analysis of the data has given the following beta (B) coefficient for the three characters

B (X1) plot size	1·00
B (X2) number of varieties	2·74
B (X3) general mean	4·18

These coefficients are obtained as a product of the partial regression coefficient and the standard deviation and measures the relative importance of the three characters in affecting the multiple correlation. Obviously, the general mean has the most effect, the number of varieties has some effect, while the plot size has the least which may very well be taken as negligible. It follows therefore that these two characters have to be taken into account to adjust the average C. V. before it is used in the equation to estimate replication number, especially when their values are far removed from their mean values. Such an adjusted value (Y) may be easily obtained from the following partial regression equation evolved from the data studied.

$$Y = 12{\cdot}57 \text{ plus } 0{\cdot}6936\ (X2 - 6) - 0{\cdot}0031\ (X3 - 2542) \quad \text{—(B).}$$

The total number of varieties to be tested is X2. X3 is the approximate value of the general mean which the trial may give and which is quite likely to be known from the preliminary trials with the varieties. 'Y' can now be calculated by substituting X2 and X3 in the above equation.

It will be noted from this equation that 'Y' varies from about 10·5 to 14·6 for the number of varieties varying from 3 to 9 if the general mean is kept constant at its mean value 2542. Similarly, 'Y' varies from about 10·8 to 14·0 for the general mean varying from 2000 to 3000, if the number of varieties under test is kept constant at 6. These two ranges of the adjusted C. V. mean fairly overlap the limits set for the mean C. V. for its random variations, viz. 12·57 ± 1·7 (see section I). It follows from these therefore, that no adjustment to the mean C. V. is quite necessary if the number of varieties and the general mean do not fall outside the limits 6 ± 3 and 2542 ± 500 respectively. If they do so it would be better if the mean C. V. is adjusted with the help of the equation (B).

Conclusion: The procedure to get the number of replications for an yield trial at the Paddy Breeding Station, Coimbatore consists of the following steps:

1. Arrive at approximate value of general mean that the experiment would give either from the records of the preliminary yield trials if available or from experience. Note down also the number of strains under test.
2. Decide upon the level of precision in terms of the difference in varietal mean yields expressed as a percentage of the general mean. This is the value of 'D' in equation (A).
3. If the general mean and the varieties under tests lie only between the limits 2500 ± 500 and 6 ± 3 respectively, adopt the value 12·57 for C. V. in equation (A).
4. If the general mean and the number of varieties fall outside the limits set in step 3, adjust the value of C. V. with the help of the equation (B).
5. Take the value of 't' in equation (A) as 1·96.
6. Substitute the values for 't' C. V. and 'D' in equation (A) and get 'n' the number of replications to be adopted in the trial.

It will be noted from the above that an average value of C. V. estimated from a representative sample of past experiments is to be first obtained for every Research station to enable use of this method for fixing the number of replications for future trials. The sample size should be as large as possible. For a majority of cases it may be adopted as such in the equation (A).

Summary: The average value of C. V. from the past 46 field experiments at the Paddy Breeding Station, Coimbatore has been worked out, its reliability tested and its use to estimate the number of replications required for future trials illustrated. Variations in number of treatment, plot size and general mean considered simulataneously accounted for only 26% of the total variability in the values of C. V. The general mean contributed the highest proportion and is thus the most important of the three factors affecting the mean C. V. A procedure to adjust the mean C. V. for variations in general mean is also suggested.

Acknowledgments: The author is much thankful to the Paddy Specialist and his staff for placing at his disposal all past records of paddy experiments at the Paddy Breeding Station, Coimbatore.

REFERENCES

Arnor, N. and Nissen, O. 1956 "Standard errors of field experiments at the Farm Crop Institute, Agric. College, Norway" *Agron. J.* **48**; 273—76.

Golden, C. H. 1952 *Method of Statistical Analysis.* Edn. 2nd John Wiley and Sons, Inc. N. Y.

Patterson D. D. 1939 *Statistical Technique in Agricultural Research.* McGraw Hill Book Publishing Company, New York.

Yield trials with paddy varieties conducted at the Paddy Breeding Station, Coimbatore, during 1953–54, 1954–55 and 1955–56

Serial No.	Coefficient of variation 'Y'	Plot size Sq. ft. 'X1'	Number of varieties 'X2'	General mean lbs./acre. 'X3'
1.	11·20	100	7	1401
2.	7.80	80	7	2394
3.	13·52	48	3	2289
4.	10·54	80	3	2970
5.	8·50	80	3	2028
6.	24·80	48	7	1585
7.	13·70	32	4	1871
8.	5·30	80	6	3048
9.	12·80	45	6	2190
10.	13·70	50	8	1868
11.	9·20	40	3	2410
12.	10·90	80	3	2788
13.	8·10	80	3	2168
14.	8·60	60	6	1908
15.	16·20	48	7	2006
16.	13.40	48	5	2294
17.	8.70	45	9	3640
18.	17·50	60	9	2604
19.	9·90	80	4	2795
20.	8.10	80	4	2036
21.	13·80	70	8	2367
22.	11·60	104	3	3674
23.	10·20	80	5	3530

Continued

Serial No.	Coefficient of variation 'Y'	Plot size Sq. ft. 'X1'	Number of varieties 'X2'	General mean lbs./acre 'X3'
24.	11·00	80	4	3742
25.	27·70	50	5	1651
26.	15·80	60	6	2425
27.	8·80	56	6	3432
28.	37·70	80	10	1554
29.	15·00	80	7	3870
30.	12·50	70	5	2881
31.	11·30	80	9	3099
32.	8·70	60	9	3519
33.	17·40	71·75	8	1997
34.	9·10	92·25	3	2851
35.	14.20	92·25	4	1274
36.	13·20	92·25	5	2486
37.	10·90	92·25	9	971
38.	9·31	92·25	5	1713
39.	11·50	92·25	4	2888
40.	17·80	92·25	3	1982
41.	9·80	92·25	5	3254
42.	12·80	70·00	9	2666
43.	6·20	82·00	6	3851
44.	6·60	71·75	6	3214
45.	12·70	92·25	4	2725
46.	11·60	71·75	9	3120
Mean	12·57	72·42	5·71	2542
Standard deviation	5·74	18·03	2·16	739·43