

# SOME CONSIDERATIONS IN DESIGNING EXPERIMENTS ON COCONUT TREES\*

BY V. J. SHRIKHANDE

*Indian Council of Agricultural Research, New Delhi*

## 1. INTRODUCTION

RESEARCH workers on coconut crop often complain that on account of large variation present between the yields of individual trees, it is not possible to derive any reliable conclusions from field experiments on cocoanut trees. The present investigation deals with certain techniques for increasing the precision of experiments on heterozygous material like coconut.

The coconut trees show a marked biennial bearing habit giving high and low yields over successive years. Since all the trees are not usually in the same phase of yield in a year, the analysis of yearly records of individual trees of any experiment may be misleading. The average yield of a coconut tree over an even number of consecutive years represents a good index of its performance and should be utilized in the analysis of data on coconut trees.

For successful experimentation with perennial plants, it is highly desirable to use biologically homogeneous material as far as possible. When experiments are performed on variable material, the use of uniformity trial records of individual trees in the pre-experimental period becomes necessary for drawing reliable conclusions.

The variability in yields of uniformly treated coconut trees is influenced mainly by (i) genetic factors and (ii) environmental factors. It will, therefore, be of interest to know, what fractions of the observed variation can be attributed to genetic and environmental factors. It is also believed that the genetic variation between trees is a more potential source of error than the environmental variation. With these considerations in view, the following problems have been investigated in this paper.

---

\* (1) Based on the author's thesis submitted in partial fulfilment of the requirements for the award of Diploma in Agricultural Statistics, Indian Council of Agricultural Research, New Delhi, 1956.

(2) The author is presently working as Senior Lecturer in Agricultural Statistics, Rafi Ahmed Kidwai Agricultural Institute, Schore, M.P.

(1) The estimation of relative contribution of genetic and environmental factors to the total variation between trees and the estimation of relative contribution of these two factors to the error variation within plots, when a tree is used as an ultimate unit of analysis.

(2) Investigating the usefulness of calibrated trees for controlling both the environmental and genetic components of error.

## 2. COMPONENTS OF VARIATION AND THEIR ESTIMATION

Pearce<sup>1</sup> has discussed this problem in connection with the determination of plot size with trees and bushes. It is well known that any phenotypic character is a joint expression of the genotype and the environment. If the genetic and environmental effects are additive and independent, the average yield  $y$  of a coconut tree over an even number of consecutive years can be expressed as  $y = g + e$ , where  $g$  is the contribution due to genotype and  $e$  is the contribution due to environment.

Let clusters of contiguous trees be formed according to a certain pattern. Let  $x$  denote the number of trees per cluster or plot, and  $c$  the total number of such clusters in the field. Let  $y_{ij}$  represent the yield of the  $j$ th tree in the  $i$ th cluster and let  $\bar{y}_i$  and  $\bar{y}$  be the means of the  $i$ th cluster and the general mean respectively. The quantities  $g_{ij}$ ,  $\bar{g}_i$ ,  $\bar{g}$ ,  $e_{ij}$ ,  $\bar{e}_i$  and  $\bar{e}$  are similarly defined for  $g$  and  $e$  components of the yield.

Using the additive model stated above, the variation between the cluster means can be expressed as follows:

$$Z = \sum_{i=1}^c \frac{(\bar{y}_i - \bar{y})^2}{c-1} = \sum_{i=1}^c \frac{(\bar{g}_i - \bar{g})^2}{c-1} + \sum_{i=1}^c \frac{(\bar{e}_i - \bar{e})^2}{c-1} + 2 \sum_{i=1}^c \frac{(\bar{g}_i - \bar{g})(\bar{e}_i - \bar{e})}{c-1}$$

If the genotypes are randomly distributed over the field, the product term on the right-hand side is expected to be zero. If the set of trees is itself a random sample of trees from some hypothetical infinite population of trees, the first quantity on the right-hand side will estimate a similar quantity about the population (*i.e.*, the same expression with summation running over population) and further if the clusters are themselves random samples of genotypes, the first quantity on the right-hand side will estimate  $\sigma_g^2/x$ , where  $\sigma_g^2$  is the variation between all the genotypes in the population and  $x$  is the size of the cluster.

The second quantity on the right-hand side estimates a quantity  $\sigma_e^2/x^b$  according to Fairfield Smith's law,<sup>2</sup> where  $\sigma_e^2$  is the variation between  $e$ 's in the population and  $b$  is a constant lying between zero and unity. Agricultural workers have verified Fairfield Smith's Law from time to time and it does generally reflect the pattern of environmental variation. Therefore  $Z$  estimates the quantity  $(\sigma_e^2/x) + \frac{\sigma_e^2}{x^b}$ .

Hence if we fit a curve of the type  $Z = (\sigma_e^2/x) + \frac{\sigma_e^2}{x^b}$  to yield data, the constants of the curve will reasonably estimate  $\sigma_e^2$ ,  $\sigma_e^2$  and  $b$  in the population. The curve can be written in the form  $xz = \sigma_e^2 + \sigma_e^2 x^{1-b}$ . If  $xz$  be denoted by  $W$ , the curve takes the form  $W = G + Ex^b$ , where the quantity  $W$  can be directly obtained as mean square between clusters in the analysis of variance table.

It will be noticed that  $x$  represents the number of trees per plot without any mention of the shape of the plot. To avoid this difficulty, all the factorisable geometric shapes are taken into account for any considered plot size superposed on the trees in a garden. Thus for  $x = 8$  the shapes  $(8 \times 1)$ ,  $(1 \times 8)$ ,  $(2 \times 4)$  and  $(4 \times 2)$  are taken where  $(8 \times 1)$  means that the plot is formed by taking 8 trees along a row,  $(4 \times 2)$  means that the plot is formed by taking 4 trees along each of the two adjacent rows and so on. Due to irregular spacing, missing trees, rejection of young trees, etc., it becomes very difficult to have all plots of the same shape for a fixed size, yet in order that lesser number of trees is wasted, sometimes some deviation in the shape of some of the plots has been allowed, although in general, the same geometric pattern has been followed as far as possible.

The size of the plot being fixed,  $W$  is calculated for each shape of plot. If there is no marked difference between these values of  $W$ , their weighted average is taken, weights being the corresponding degrees of freedom. The curve then takes the form  $W' = G + Ex^b$ , where  $W'$  is the weighted average.

*A simple method of curve fitting.*—The curve can also be written in the form  $W' = G + Ee^{B \log x}$ , which resembles the form  $y = \alpha + \beta \rho^x$ . Several workers have discussed the fitting of this curve,<sup>3-6</sup> where in all such cases except in<sup>(3)</sup>,  $x$  is usually taken to be in arithmetic progression, which is not true of  $\log x$ . Here a simple device has been used to fit this curve. It takes advantage of the knowledge that  $B$  is expected to lie between zero and one. For any trial value of  $B$  between zero and one,  $x^B$  is calculated for different values of  $x$ . If  $x^B = x_1$

where  $x_1$  is now known, the problem reduces to fitting a straight line  $W' = G + Ex_1$  and estimating  $B$  by a quantity which gives the minimum residual  $S.S.$  Thus by trial and error, the requisite value  $B_0$  is found out, correct up to two places of decimal. The method has been found to be successful in some of the practical applications though it does not provide directly the estimate of the variance of  $B_0$ . The standard errors of  $G$  and  $E$  can be calculated approximately by assuming  $x^{20}$  to be an errorless variable.

If  $G_0$ ,  $E_0$  and  $B_0$  are the calculated constants of the curve and  $(R.S.S.)_0$  is the corresponding minimum residual  $S.S.$ , the approximate standard errors of  $G_0$  and  $E_0$  are given by

$$S.E. (G_0) = \sqrt{\sigma_0^2 \left[ \frac{1}{n} + \frac{\bar{x}_1^2}{S(x_1^2)} \right]}$$

$$S.E. (E_0) = \sqrt{\sigma_0^2 \frac{1}{S(x_1^2)}}$$

where

$$\sigma_0^2 = \frac{(R.S.S.)_0}{n-3}, \quad S(x_1^2) = \sum (x_1 - \bar{x}_1)^2$$

and  $n$  is the number of points. It will be noticed that for  $x=1$ ,  $W' = W$  the total variation between trees estimates  $\sigma_g^2 + \sigma_e^2$ . Hence  $G_0$  and  $E_0$  are the estimates of genetic and environmental components of the total variation between trees.

*Genetic and environmental components of within cluster mean square.*—We shall now investigate into the contribution of genetic and environmental components of the error variation within clusters of trees, a tree being an ultimate unit of analysis. Using the same model  $y = g + e$  for a tree we have,

$$\begin{aligned} \sum_{i=1}^c \sum_{j=1}^a \frac{(y_{ij} - \bar{y}_i)^2}{c(x-1)} &= \frac{\sum \sum (g_{ij} - \bar{g}_i)^2}{c(x-1)} + \frac{\sum \sum (e_{ij} - \bar{e}_i)^2}{c(x-1)} \\ &+ \frac{2\sum \sum (g_{ij} - \bar{g}_i)(e_{ij} - \bar{e}_i)}{c(x-1)} \end{aligned}$$

If we assume that the genotypes are randomly distributed over the field, the product term on the right-hand side may be neglected. Again if each cluster is a random sample of genotypes, the genetic variation within each cluster is an estimate of the genetic variation in the infinite population; hence the average of the genetic variation within each

cluster, over different clusters is also an estimate of the genetic variation in the population. In other words

$$\frac{\Sigma\Sigma (g_{ij} - \bar{g}_i.)^2}{c(x-1)}$$

is an estimate of  $\sigma_g^2$ .

To calculate the expectation of the second term on the right-hand side, we make use of the following identity.

$$\sum_{i=1}^c \sum_{j=1}^a (e_{ij} - \bar{e})^2 = \Sigma\Sigma (e_{ij} - \bar{e}_i.)^2 + x \sum_{i=1}^c (\bar{e}_i. - \bar{e})^2.$$

If

$$\frac{\Sigma\Sigma (e_{ij} - \bar{e})^2}{(cx-1)} = s_e^2$$

and

$$\frac{\Sigma\Sigma (e_{ij} - \bar{e}_i.)^2}{c(x-1)} = s_{ew}^2$$

then the identity can be written in the form

$$(cx-1)s_e^2 = c(x-1)s_{ew}^2 + x(c-1) \sum_{i=1}^c \frac{(\bar{e}_i. - \bar{e})^2}{(c-1)}.$$

Taking expectations on both sides and using Fairfield Smith's Law, the identity becomes

$$(cx-1)\sigma_e^2 = c(x-1)\sigma_{ew}^2 + x(c-1)\frac{\sigma_e^2}{x^b}$$

or

$$\sigma_{ew}^2 = \frac{\left(x - \frac{1}{c}\right)\sigma_e^2 - x\frac{(c-1)\sigma_e^2}{x^b}}{(x-1)}.$$

For a large number of clusters we have in the limit

$$\sigma_{ew}^2 = \frac{x}{x-1} (1-x^{-b})\sigma_e^2$$

which implies that in the limit  $s_{ew}^2$  estimates  $\frac{x}{x-1} (1-x^{-b})\sigma_e^2$ .

Therefore the analysis of variance table with the expectation of mean squares can be presented as given in the next page.

TABLE I.

| Source of Variation | <i>d.f.</i> | <i>M.S.</i>   | Expectation <i>M.S.</i>                              |
|---------------------|-------------|---|--|
| Between clusters    | $(c - 1)$   | $x \sum_{i=1}^c \frac{(\bar{y}_i - \bar{y})^2}{c - 1}$            | $\sigma_g^2 + x \frac{\sigma_e^2}{x^b}$              |
| Within clusters     | $c(x-1)$    | $\sum_{i=1}^c \sum_{j=1}^x \frac{(y_{ij} - \bar{y}_i)^2}{c(x-1)}$ | $\sigma_g^2 + \frac{x}{x-1} (1 - x^{-b}) \sigma_e^2$ |
| Total               | $(cx - 1)$  | $\frac{\sum \sum (y_{ij} - \bar{y})^2}{(cx - 1)}$                 | $\sigma_g^2 + \sigma_e^2$                            |

The expectation of mean squares is true when  $c$  tends to infinity.

The multiplying factors of  $\sigma_e^2$  in Table I are  $x/x^b$  and  $(x/x-1)(1-x^{-b})$ . The quantity  $x/x^b$  is expected to be greater than 1 as 'b' is expected to lie between zero and one. For this reason between cluster *M.S.* is expected to increase with  $x$ , whereas the quantity

$$\frac{x}{x-1} (1 - x^{-b}) = \frac{\left(1 - \frac{1}{x^b}\right)}{\left(1 - \frac{1}{x}\right)}$$

is positive and clearly less than 1. This quantity slowly increases with  $x$  and tends to unity as  $x$  becomes very large. Therefore the *M.S.* within clusters is a function which increases very slowly with  $x$ . The sampling fluctuations in the estimates of  $\sigma_g^2$  and  $\sigma_e^2$  associated with any particular type of grouping, may mask the slowly rising trend in the Within cluster *M.S.* It is, therefore, better to estimate  $\sigma_g^2$  and  $\sigma_e^2$  using between cluster *M.S.*, which is expected to increase more rapidly with  $x$  than the Within cluster *M.S.*

The estimates of  $\sigma_g^2$  and  $\sigma_e^2$  may now be used to obtain the estimate of Within cluster *M.S.* As a check on the goodness of fit of these estimates, we might compare the expected value of Within cluster *M.S.* with the weighted average of the actual values of Within cluster *M.S.* over different shapes for a given plot size  $x$ .

The above findings are tested on the uniformity data obtained from Block F of the Coconut Research Station, Pilicode and Field IX of the Central Coconut Research Station, Kasargod. The results are presented in Section 3.

*Limitations of the method.*—The method depends upon the validity of the following assumptions:

- (i) Fairfield Smith's Law holds good.
- (ii) The Phenotype can be split into two additive parts—genetic and environmental.
- (iii) The genotypes are randomly distributed over the field and
- (iv) Each cluster is a random sample of genotypes.

The curve fitted is

$$Z = \frac{\sigma_g^2}{x} + \frac{\sigma_e^2}{x^b}$$

The constant 'b' is associated with the intraclass correlation between  $e$ 's within plots. The value of 'b' depends to some extent on the nature of fertility contours in the field. It may sometimes happen that  $b$  may be unity or not much different from unity, in which case the curve becomes  $xZ = \sigma_g^2 + \sigma_e^2$ , in which case the separation of the two components of variation may not be possible.

Subject to these limitations, the method does give an approximate idea of the genetic and environmental components of the total variation between trees and their contribution to the Within cluster variation for different plot sizes.

It appears from the following illustrations that with coconut crop, the genetic component is greater than the environmental component in the total variation between trees. Even if these components are equal, the genetic component of error variation within clusters is much larger than the corresponding environmental component. This is because, the multiplying factor of  $\sigma_e^2$  in the expectation of the Within cluster *M.S.* is  $x/x - 1 [1 - x^{-b}]$  which is usually a very small positive quantity less than unity. This quantity increases very slowly with  $x$  and in the limit tends to unity as  $x$  tends to infinity.

If the genetic component of error variation is higher than the environmental component, alternative methods must be devised which aim at controlling the genetic component using the same set of trees. This is discussed in Section 4.

### 3. PRACTICAL APPLICATIONS

*Example 1.*—Between cluster and Within cluster mean squares for 172 trees of Block F at Pilicode are presented in Table-II.

TABLE II  
Between cluster and Within cluster Mean Squares

| Plot Size (x) | Shape                      | Between cluster M.S. (W) | d.f. | Pooled Between cluster M.S. (W') | Within cluster M.S. | d.f. | Pooled Within cluster M.S. |
|---------------|----------------------------|--------------------------|------|----------------------------------|---------------------|------|----------------------------|
| 1             | ..                         | 191.49                   | 171  | 191                              | ..                  | ..   | ..                         |
| 2             | { 2×1<br>1×2               | 260.67                   | 84   | 232                              | 123.53              | 85   | 151.33                     |
|               |                            | 204.33                   | 85   |                                  | 178.80              | 86   |                            |
| 3             | { 3×1<br>1×3               | 261.85                   | 56   | 270                              | 140.16              | 114  | 144.82                     |
|               |                            | 279.17                   | 56   |                                  | 149.49              | 114  |                            |
| 4             | { 4×1<br>1×4<br>2×2        | 311.40                   | 41   | 282                              | 154.56              | 126  | 159.83                     |
|               |                            | 251.30                   | 42   |                                  | 172.02              | 129  |                            |
|               |                            | 283.39                   | 41   |                                  | 152.63              | 126  |                            |
| 5             | { 5×1<br>1×5               | 384.59                   | 33   | 347                              | 145.67              | 136  | 155.58                     |
|               |                            | 308.96                   | 33   |                                  | 165.49              | 136  |                            |
| 6             | { 6×1<br>1×6<br>2×3<br>3×2 | 335.02                   | 27   | 394                              | 164.47              | 140  | 152.87                     |
|               |                            | 376.40                   | 27   |                                  | 157.94              | 140  |                            |
|               |                            | 391.20                   | 27   |                                  | 151.25              | 140  |                            |
|               |                            | 471.95                   | 27   |                                  | 137.85              | 140  |                            |
| 8             | { 8×1<br>1×8<br>4×2<br>2×4 | 417.68                   | 20   | 415                              | 164.58              | 147  | 159.05                     |
|               |                            | 417.61                   | 20   |                                  | 163.14              | 147  |                            |
|               |                            | 351.55                   | 19   |                                  | 153.13              | 140  |                            |
|               |                            | 470.32                   | 20   |                                  | 155.06              | 147  |                            |

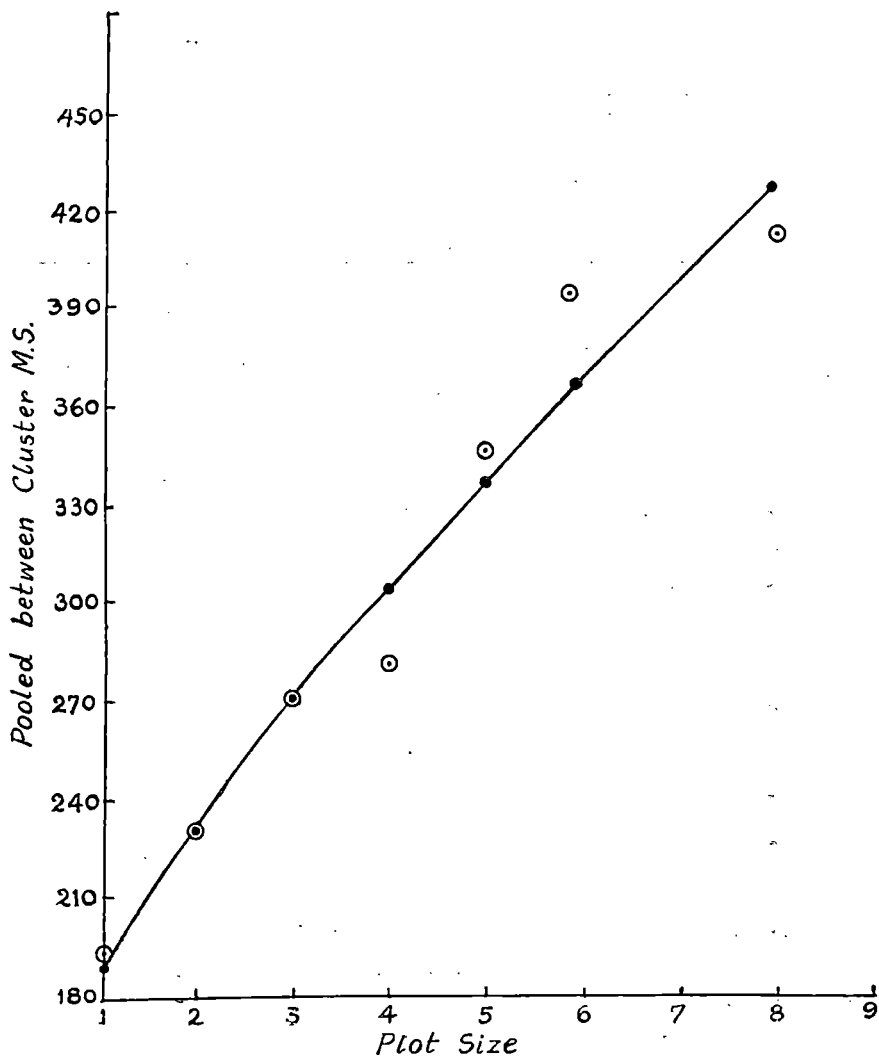
Here  $W'$  shows a clearly rising trend (Graph I) with increase in  $x$ , whereas pooled Within cluster  $M.S.$  shows only a very slightly rising trend with  $x$ .

For the sake of illustration, the trial values of  $B$  and the values of the corresponding residual sum of squares are presented below:

TABLE III

| $B$           | ..      | 0.50    | 0.60    | 0.70    | 0.74    | 0.75    | 0.76    | 0.80    | 1.00 |
|---------------|---------|---------|---------|---------|---------|---------|---------|---------|------|
| Residual S.S. | 1630.44 | 1477.05 | 1410.40 | 1407.30 | 1407.09 | 1409.33 | 1424.12 | 1673.67 |      |





GRAPH I (Block F, Pilicode)

Graph showing rising trend of  $W'$  and Agreement between  $W'$  observed and  $W'$  expected.

The curve fitted between  $W'$  and  $x$  has the equation,

$$W' = 124.7900 + 63.7536 x^{0.75}$$

The curve accounts for 96.64% of the variation in  $W'$ . The observed and the expected values of  $W'$  are presented in Table IV.

TABLE IV  
*Comparison of Observed and Expected Values of  $W'$*

| Plot Size | $W'$ (observed) | $W'$ (expected) |
|-----------|-----------------|-----------------|
| 1         | 191             | 188.54          |
| 2         | 232             | 232.01          |
| 3         | 270             | 270.12          |
| 4         | 282             | 305.11          |
| 5         | 347             | 337.96          |
| 6         | 394             | 369.20          |
| 8         | 415             | 428.05          |

There is a close agreement between the observed and the expected values of  $W'$ . The estimates of  $\sigma_p^2$  and  $\sigma_e^2$  are

$$G_0 = 124.7900 \quad S.E. (G_0) = 17.960$$

$$E_0 = 63.7536 \quad S.E. (E_0) = 5.945$$

$$B_0 = 0.75$$

The genetic and environmental components of the total variation between trees are approximately in the ratio of 2:1. The estimate of  $b$  is 0.25.

The environmental component of the within cluster variation is estimated by  $(\hat{E} = x/x - 1) (1 - x^{-0.25}) E_0$  and the expected value of Within cluster  $M.S.$  is estimated by  $(G_0 + \hat{E})$ . These two quantities along with the average values of Within cluster  $M.S.$  corresponding to different plot sizes are presented below.

TABLE V

| $x$ | $\hat{E}$ | $G_0 + \hat{E}$ | Average Within cluster $M.S.$ |
|-----|-----------|-----------------|-------------------------------|
| 2   | 20.29     | 145.08          | 151.33                        |
| 3   | 22.97     | 147.76          | 144.82                        |
| 4   | 24.90     | 149.69          | 159.83                        |
| 5   | 26.40     | 151.19          | 155.58                        |
| 6   | 27.62     | 152.41          | 152.87                        |
| 8   | 29.52     | 154.31          | 159.05                        |

The agreement between the observed and the expected values of Within cluster *M.S.* appears to be fairly good. The ratio of genetic to environmental components of the within cluster variation varies between 6:1 and 4:1 approximately for different plot sizes.

*Example 2.*—Between cluster and Within cluster mean squares for 168 trees at Kasargod are presented below:

TABLE VI  
*Between cluster and Within cluster Mean Squares*  
*(Field IX Kasargod)*

| Plot Size ( <i>x</i> ) | Shape | Between cluster <i>M.S.</i> ( <i>W</i> ) | <i>d.f.</i> | Pooled Between cluster <i>M.S.</i> ( <i>W'</i> ) | Within cluster <i>M.S.</i> | <i>d.f.</i> | Pooled Within cluster <i>M.S.</i> |
|------------------------|-------|--|-------------|--|----------------------------|-------------|-----------------------------------|
| 1                      | ..    | 349.33                                   | 167         | 349  | ..                         | ..          | ..                                |
| 2                      | 2×1   | 369.33                                   | 80          | 386  | 282.33                     | 81          | 293.48                            |
|                        | 1×2   | 402.91                                   | 80          |  | 304.63                     | 81          |                                   |
| 4                      | 1×4   | 500.16                                   | 40          | 576  | 293.21                     | 123         | 273.96                            |
|                        | 4×1   | 585.98                                   | 40          |  | 273.31                     | 123         |                                   |
|                        | 2×2   | 642.37                                   | 40          |  | 255.38                     | 123         |                                   |
| 6                      | 1×6   | 541.33                                   | 27          | 653  | 312.30                     | 140         | 288.74                            |
|                        | 6×1   | 640.32                                   | 26          |  | 284.06                     | 135         |                                   |
|                        | 2×3   | 816.83                                   | 26          |  | 266.29                     | 135         |                                   |
|                        | 3×2   | 619.84                                   | 26          |  | 291.44                     | 135         |                                   |
| 8                      | 4×2   | 452.71                                   | 18          | 746  | 242.87                     | 133         | 273.96                            |
|                        | 2×4   | 828.87                                   | 20          |  | 284.09                     | 147         |                                   |
|                        | 1×8   | 792.93                                   | 20          |  | 288.98                     | 147         |                                   |
|                        | 8×1   | 881.45                                   | 20          |  | 276.93                     | 147         |                                   |
| 10                     | 1×10  | 727.51                                   | 15          | 830  | 293.74                     | 144         | 297.26                            |
|                        | 10×1  | 758.58                                   | 15          |  | 318.82                     | 144         |                                   |
|                        | 2×5   | 947.24                                   | 15          |  | 285.80                     | 144         |                                   |
|                        | 5×2   | 887.51                                   | 15          |  | 290.69                     | 144         |                                   |

The curve fitted between *W'* and *x* has the equation

$$W' = 172.7627 + 161.8401 x^{0.61}$$

The curve accounts for 98.90% of the variation in *W'*. The observed and the expected values of *W'* from the curve are presented in Table VII.

TABLE VII  
*Comparison of Observed and Expected Values of  $W'$*

| $x$ | $W'$ (observed) | $W'$ (expected) |
|-----|-----------------|-----------------|
| 1   | 349             | 334.60          |
| 2   | 386             | 419.78          |
| 4   | 576             | 549.78          |
| 6   | 653             | 655.55          |
| 8   | 746             | 748.22          |
| 10  | 830             | 832.07          |

The rising trend of  $W'$  with increasing  $x$  and the agreement between  $W'$  observed and  $W'$  expected are shown in Graph II. The estimates of  $\sigma_g^2$ ,  $\sigma_e^2$  and their standard errors are given below:

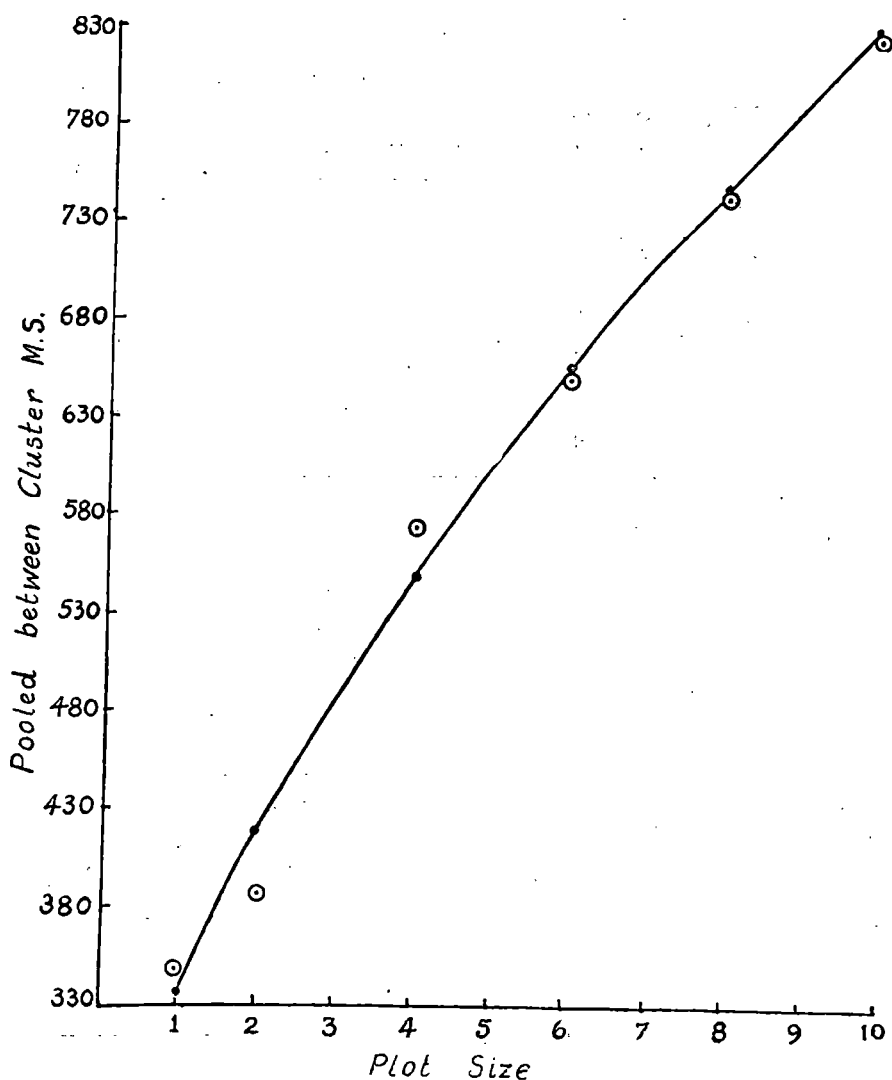
$$\begin{aligned}
 G_0 &= 172.7627 & S.E. (G_0) &= 27.6030 \\
 E_0 &= 161.8401 & S.E. (E_0) &= 9.8700 \\
 B_0 &= 0.61
 \end{aligned}$$

The estimate of  $b$  is 0.39, which is less than unity as expected. In this case the genetic and environmental components of the total variation between trees are approximately of the same magnitude.

The following table gives the environmental components of the error variation within clusters and the observed and expected values of the Within cluster *M.S.*

TABLE VIII

| $x$ | $\hat{E}$ | $G_0 + \hat{E}$ | Average Within cluster <i>M.S.</i> |
|-----|-----------|-----------------|------------------------------------|
| 2   | 76.68     | 249.44          | 293.48                             |
| 4   | 90.13     | 262.89          | 273.96                             |
| 6   | 97.65     | 270.42          | 288.74                             |
| 8   | 102.78    | 275.55          | 273.95                             |
| 10  | 106.56    | 279.32          | 297.26                             |



GRAPH II (Field IX, Kasargod)

Graph showing rising trend of  $W'$  and Agreement between  $W'$  observed and  $W'$  expected.

In this case the ratio of genetic to environmental components of the within cluster variation varies between 2:1 to 3:2 approximately, for different plot sizes.

The data for one more field at Kasargod was examined and it was found that  $G:E = 2:1$  approximately. Also the ratio of genetic to

environmental components in the within cluster variation varied approximately between 10:1 and 6:1 for different plot sizes.

#### 4. ALTERNATIVE METHODS OF CONTROLLING ERROR VARIATION

It appears that the genetic component in the expression for the yield of a tree is more important than the corresponding environmental component. Also the variation between  $g$  components of trees seems to be more than that between the corresponding  $e$  components.

The usual method of lay-out of an experiment is to divide the land into compact blocks and within each block the adjacent trees are grouped together to form plots. Let this be called Method I. This method aims at reducing within block variation and increasing the between block variation, as far as the  $e$  component of yield is concerned.

Since the variation between  $g$  components is expected to be larger than that between  $e$  components, a method which groups together similar  $g$  components within a block, should be more successful in controlling the error variation. It is expected that the correlation between  $y$  and  $g$  should be higher than that between  $y$  and  $e$ . This suggests an alternative method as given below.

Let the trees be arranged in descending order of magnitude according to their yield performance, *i.e.*, the total crop of a tree over even number of consecutive years. Suppose there are four treatments to be tried in 8 tree plots. The ordered trees are divided into groups of 32 trees. Each group of 32 trees may be called a block, where the block is no more a compact piece of land but a group of relatively homogeneous genotypes. In this block of ordered trees, apply the treatments  $A, B, C, D$  at random to the first four trees, then to the next four trees and so on, till all the trees in that block are exhausted. In this block, the 8 trees to which treatment  $A$  is applied form a plot, all the 8 trees to which treatment  $B$  is applied form another plot and so on. Thus there are four plots in this block. Similarly the other blocks are dealt with. Now if the model of a randomised block design is assumed for a plot, the analysis can be performed as usual and the coefficient of variation ( $C.V.$ ) corresponding to plot error can be calculated. Let this be called Method II. This method aims at reducing within block variation and increasing the between blocks variation as far as  $g$  component is concerned.

The two methods can be combined into a third method, which is as follows. First divide the land into compact blocks of 32 trees. The trees within blocks are arranged according to the total crop production

per tree. The same method of randomisation of treatments within a block is followed as in Method II. In a block, the trees having the same treatment form a plot. Let this be called Method III. This method aims at reducing the within block variation, by making plots within compact blocks as homogeneous as possible for  $g$  components while compact blocks are used to control the environmental variation.

In Methods II and III it is necessary that the blocks and plots should be decided upon, by the use of total crop production per tree in a given period and the data for subsequent period should be analysed for these blocks and plots. This should be so, because in practice plans for blocks and plots would be prepared on the basis of pre-experimental uniformity records and the data for these blocks and plots would be analysed for the subsequent experimental period.

For different plot sizes, these methods have been studied on the records at Pilicode and Kasargod. With each of the three methods, the yield of a plot in the pre-experimental period is treated as a concomitant variate and the analysis of covariance has been performed. The results of such analysis are given below, the design used being a Randomised Block Design with 4 hypothetical treatments.

TABLE IX

*Block F of Pilicode*

$y$  = Total plot yield for 1953 and 1954.

$x$  = Total plot yield for 1949 to 1952 for the same plot.

*C.V. corresponding to plot error and plot mean*

| Plot Size | Method I | Method I with co-variance | Method II | Method II with co-variance | Method III | Method III with co-variance |
|-----------|----------|---------------------------|-----------|----------------------------|------------|-----------------------------|
| 1         | 35.22    | 26.82                     | 24.27     | 24.36                      | 35.22      | 26.82                       |
| 2         | 27.76    | 20.53                     | 16.94     | 16.91                      | 22.20      | 19.01                       |
| 4         | 25.26    | 15.54                     | 9.96      | 9.99                       | 10.54      | 10.05                       |
| 6         | 15.73    | 11.93                     | 9.24      | 8.78                       | 13.43      | 11.07                       |
| 8         | 14.58    | 7.64                      | 8.99      | 8.99                       | 12.04      | 9.89                        |

TABLE X

*Field I Kasargod*

$y$  = Total plot yield for 1951 and 1952.

$x$  = Total plot yield for 1949 and 1950 for the same plot.

*C. V. corresponding to plot error and plot mean*

| Plot Size | Method I | Method I with co-variance | Method II | Method II with co-variance | Method III | Method III with co-variance |
|-----------|----------|---------------------------|-----------|----------------------------|------------|-----------------------------|
| 1         | 37.13    | 19.05                     | 20.27     | 20.26                      | 37.13      | 19.05                       |
| 4         | 26.05    | 12.37                     | 10.12     | 10.19                      | 12.07      | 8.80                        |
| 8         | 21.51    | 10.21                     | 7.06      | 7.07                       | 9.15       | 9.41                        |
| 12        | 18.11    | 5.75                      | 6.94      | 6.06                       | 3.98       | 3.77                        |

It is clearly brought out that the Methods II and III which aim at controlling genetic variation, are more effective than Method I which aims to control the environmental variation.

Methods II and III make use of the previous records to arrange blocks and plots whereas the first method does not make use of the past records for such planning. For this reason, Method I with covariance is comparable with Method II or Method III. Though more investigations are necessary to decide the superiority of one method over the other, there is some evidence to show that Method II is slightly better than Method I with covariance. In one of the fields Method III gives slightly better performance than Method I with covariance, though for the other field it is not so.

Method III with covariance scores over Method I with covariance. It is observed that the use of covariance technique with Methods II or III does not reveal a substantial improvement, which is expected, because covariance with Methods II or III means a two-fold use of the previous records, viz., (i) To plan blocks and plots, (ii) To use past records for covariance correction. It is obvious that once the past records are used, not much information can be extracted out of them again.



It should be noted that for one tree plots Method I and Method III are identical.

If the treatments to be applied are such that when a treatment is applied to a tree the neighbouring trees are not affected, the last two methods may be recommended. If on the grounds of practical convenience or cost considerations these methods are unsuitable, it is advisable to use Method I with covariance.

#### SUMMARY

It has been suspected for a long time that the genetic variation between cocoanut trees is a more potential source of error than the environmental variation. An attempt has been made to estimate the relative contribution of genetic and environmental factors to the total variation between trees, by modifying the well-known Fairfield Smith's empirical law describing heterogeneity in the yields of agricultural crops, to include variation due to genetic factors. Having determined these relative contributions, the components of within plot variation due to genetic and environmental factors have been worked out, for different plot sizes, a tree being an ultimate unit of analysis. Alternative methods of field plot technique through grouping of trees having similar yields have been investigated.

#### ACKNOWLEDGEMENTS

The author is very much indebted to Dr. Uttamchand, Professor of Statistics, Indian Council of Agricultural Research, New Delhi, under whose guidance the investigation was carried out. He would also like to thank Dr. V. G. Panse, Statistical Adviser, Indian Council of Agricultural Research, New Delhi, for his valuable suggestions and guidance from time to time during the course of investigation. The uniformity data used in this paper was obtained through the courtesy of the Director, Central Coconut Research Station, Kasargod. The permission to use the data is gratefully acknowledged.

#### REFERENCES

1. Pearce, S. C. . . "Some considerations in deciding plot size in field trials with trees and bushes," *J. Ind. Soc. Agri. Stat.*, 1955, 7.
2. Smith, H. Fairfield . . An empirical law describing heterogeneity in the yield of agricultural crops," *J. Agri. Sci.*, 1938, 28, 1-23.

3. Deming, W. E. .. *Statistical Adjustment of Data*, John Wiley & Sons, New York, 1948.
4. Gomes, F. P. .. "The use of Mitscherlich's regression law in the analysis of experiments with fertilizers," *Biometrics*, 1953, 9 (4), 498-516.
5. Hartley, H. O. .. "The estimation of nonlinear parameters by internal least squares," *Biometrika*, 1948, 35, 32-45.
6. Stevens, W. L. .. "Asymptotic Regression," *Biometrics*, 1951, 7, 247-67.