3

Problems of Experimentation with Cocoa Trees

By R. LOTODÉ and RAOUL A. MULLER

Part I. The necessity of field experimentation with cocoa: difficulties encountered

Need for work in the plantation

Experiments aimed at increasing the yield of the cocoa tree (cultivation techniques, methods of pest control) must be made in the conditions where they would normally be applied: results of experiments must be obtained in ecological situations comparable to those where they will be put into practice.

In most cases there are too few agronomic experiment stations to be representative of all the ecological zones that exist in a producing country. Most experimental work, and particularly that dealing with plant health problems, must therefore be done in plantations.

Characteristics of cocoa plantations

Very generally, cocoa plantations are characterized by extreme heterogeneity: this heterogeneity is of several kinds.

Genetic heterogeneity
It is very rare to find plantations consisting of a single clone. More often the trees are grown from seed and are more or less hybrid. The resulting diversity is particularly large where, as in Cameroon, Trinitario are cultivated originating from crosses between Criollo and Forastero: a single plantation may thus contain trees differing greatly in habit of growth, vegetative behaviour, fruit characteristics, and potential yield.

Heterogeneity of terrain
In new countries cocoa is usually planted on recently cleared land, whose soil has never been homogenized either in physical structure or chemical composition, as has soil which has been cultivated over a long period. In addition, cocoa

23

is usually cultivated under natural shade, the various species of which have different requirements and which compete differently with the cocoa trees for water and minerals.

Microclimatic heterogeneity

Because the cocoa tree is a shade-loving plant, plantations are generally started in the remains of forest resulting from more or less intensive thinning of the forest, the choice of which trees to keep and which to remove being partly dictated by ease of felling, which depends on the means (generally inadequate) available for this operation. These vestiges of forest thus form a very irregular cover, where very dense, shady, poorly ventilated and moist areas can be found side by side contrasting with areas that are much more open, sunny and consequently drier. The behaviour of the trees, their growth and their fruiting capacity are very diversely affected by the cover. The activity of parasites, insects (mirids in particular) and fungi (above all *Phytophthora palmivora*) under the influence of environmental factors, all show considerable variations within one and the same plantation.

Heterogeneity due to age

It is noteworthy that trees of different ages are found intermingled in the cocoa plantation as a result of replacements over a period of time following various

Table 1   Variations under natural conditions: (1) of mean yield of pods per tree; (2) of damage due to black pod

Year of observation	Plantation	Block	Plot	Mean yield of pods per tree	Percentage of losses due to black pod
1957–8	1	1	1	17·2	18·0
			2	29·2	31·7
			3	20·7	12·5
			4	24·1	46·6
		2	1	25·3	39·7
			2	13·2	21·2
			3	29·0	30·3
			4	36·4	40·6
	2	1	1	15·1	66·1
			2	19·2	57·5
			3	36·1	51·0
			4	23·1	67·2
	3	1	1	27·5	74·5
			2	46·8	74·9
			3	38·6	73·7
			4	34·4	62·0
		2	1	46·5	69·3
			2	27·8	49·5
			3	25·6	46·2
			4	22·4	46·0

accidents happening to individual trees. Furthermore, when the plantations are old, as happens for example in Brazil, the trees take on tortuous and highly varied shapes which no doubt affect their productivity and chances of being parasitized.

This multiple heterogeneity of the cocoa tree results in large variations in yield and parasitic attack between adjacent plots. By way of illustration, Table 1 gives figures for yield and losses due to black pod recorded in Cameroon from adjacent plots, each comprising about 200 trees and marked out in apparently homogeneous plantations.

This table shows in particular that between two adjacent plots in the same block:

yield of cocoa pods can vary in a proportion ranging from 1 to 3;
losses due to black pod can vary in a proportion ranging from 1 to 4;
a regular gradient of variation can appear within a single plantation of approximately 2 hectares in respect to the incidence of black pod.

We were able to observe in Brazil (and this is evident from the Annual Report for 1968 and 1969, CEPEC Phytopathology Laboratory, Itabuna, Bahia) that variations can be found comparable with those described in Cameroon, both in respect to yield and the incidence of black pod on adjacent experimental plots.

Difficulties involved in field experimentation

This heterogeneity of cocoa plantations and the marked variations in agricultural and parasitic phenomena resulting from it run counter to the rules of agricultural experimentation, which in order to be brought to a successful conclusion must be carried out in an environment as homogeneous as possible: it must be possible to distinguish even quite small effects of applied treatments from natural random variations.

To illustrate this we shall cite an insurance from Cameroon between 1955 and 1956, where a group of 500 apparently homogeneous trees were kept under observation by the Phytopathology Laboratory of the Agronomic Research Centre of Nkolbisson and divided into 5 blocks, each containing 10 plots of 10 trees. No treatment was given, and observations were made plot by plot only. An analysis of this uniformity trial, based on Fisher's method of blocks, should, of course, have produced identical statistics for the several recorded yields. It did nothing of the kind: the average crop from the plots in each block were very significantly different, and moreover arrangements of 5 plots taken at random, one from each block, did not themselves have comparable means, once the variability due to the block effect had been eliminated.

It may be objected that in this trial the plots of 10 trees were too small, bringing out the genetic heterogeneity of the trees, and that 5 replicates were insufficient.

Another trial using Fisher's block design, with the aim of comparing sufficiently large plots (approximately 200 trees) in order to eliminate the effect of genetic heterogeneity of the trees, did not give any more satisfactory results.

This trial comprised, in 1958–9, 5 blocks of 4 plots, corresponding to three treatments against black pod and one control. It was repeated in 1959–60, and during this second year a second series of 4 identical blocks was started (giving a total of 9 replicates).

Analysis of the results by the conventional method (comparison of the means of groups of plots receiving the different treatments) was very disappointing. Without going into details, suffice it to say that in 1958–9 differences of the order of 30 per cent in yield of healthy pods, and for 1959–60 differences of the order of 1 to 8 in percentage of black pods, did not show as significant. Under these conditions it is quite obvious that any attempts at experimentation are useless.

While it may be assumed that in this trial the large size of the plots led to the partial elimination of variations due to genetic heterogeneity of the trees, it is to be feared, on the other hand, that the effects of heterogeneity of soil and microclimate were amplified.

We in Cameroon are not the only ones to have experienced such failures due to the very nature of the environment: the Annual Report of the CEPEC Phytopathology Laboratory for the year 1968–9 brings out the fact that the Brazilian investigators are faced with similar problems. Furthermore, the specialist literature often reports experiments with cocoa trees which remained inconclusive or have led to only vague conclusions.

It may be recalled at this point that the Inter-American Conference on Cacao (5th and 7th meetings, held at Turrialba, Costa Rica, in 1954 and Palmira, Colombia, in 1958 respectively) expressed concern about this very general problem and recommended the compulsory definition of experimental designs adapted to conditions in cocoa plantations.

Of course it is possible that in the foreseeable future some of the difficulties now encountered in the plantation will become less acute with the use of genetically homogeneous cultivars and more highly developed methods of cultivation. Particular attention should be paid to the present general tendency towards cultivation without shade, or with reduced shade, thus avoiding or decreasing the hazards due to forest cover. But if the trend in cocoa growing is taking the general direction of a simplification of the problems facing the experimenter, it will not however eliminate all the difficulties. It is still true that existing cocoa plantations, however poorly adapted for experimental work, will remain for a long time to come the main area of experimental material available to research workers.

In view of the pressing need for field experimentation, with all the difficulties which this involves, certain aspects and causes of which have just been analysed, research is essential in order to delve more deeply into certain problems relating to experimental designs and techniques for analysing the results.

It will be necessary:

1. to reduce as much as possible causes of variability, other than those whose effects are to be measured;
2. to determine the optimum size of the elementary plot in order to obtain the greatest possible precision;

3. to include in the analysis of variance those factors whose effect can be measured and which, if significant, will result in the reduction of the residual error and therefore ensure greater precision (analysis of covariance);
4. to ensure the validity of the analysis of variance by appropriate transformation of the data;
5. to choose the criterion for judging the effect of an applied treatment which is best adapted for the end in view.

Very little methodological work specially adapted to the problems of the cocoa tree has yet been done in the world. Accordingly we have been obliged to rely on our own experiments in Cameroon in trying to contribute to the clarification of this problem.

Part II. Aspects of experimentation designed to compare yields from contrasting treatments or varieties

Structural study of yield, and definition of optimum elementary plot size

Structure of yield

This is the most general and frequently occurring case in experiments whose purpose is to gain knowledge either of planting material, of growing techniques, or of the effects on yields of applying plant health treatments integrated with current agricultural practices.

Yield as a criterion for estimating the effect of an applied treatment has been the subject of two studies made with very different groups of trees, one (No. 1) being a collection whose individuals had been studied for ten years at the Nkoemvone (South Cameroon) cocoa station, and the other (No. 2) being made up of various traditional type plantations in the Yaounde region (Central Cameroon).

It is noteworthy that these studies, although made in very different conditions, have led to essentially similar conclusions.
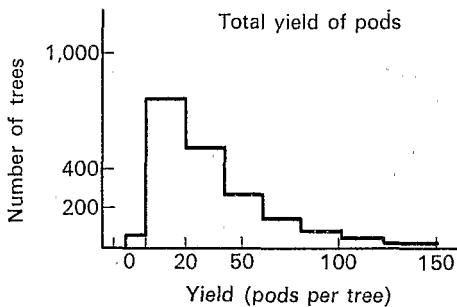

Fig. 3 Asymmetric distribution of yields per tree in traditional plantations in the Yaounde region.

The structural analysis of the yields of the individual trees in traditional plantations (No. 2) showed that the distribution is not normal but highly skew, as shown by the histogram (Fig. 3). Similarly the variances within plots are not homogeneous.

A similar study was made with cumulative individual yields of trees in the Nkoemvone Station collection (No. 1), ten years after planting. This covers data from 5 fields (plots 21 to 25 of the Nkoemvone collection). In each field elementary plots comprising: 1, 4, 8, 12, 16, 20, 24, 32, 36, 40, 64, 96 trees were marked off. The plots were set up in the most compact manner possible. They were not entirely regular in shape owing to missing individuals. The variances of the raw means of the plots, and after transformation to logarithms (base 10), are given in Table 2.

Table 2

Size of Plots	Field I – Gen. Mean: 4,925 (576 trees)			Field II – Gen. Mean: 5,705 (420 trees)		
	Variance raw data*	Variance transformed data	d.f.	Variance raw data	Variance transformed data	d.f.
1	13·452	0,18091	575	27·679	0,16187	419
4	8·009	0,09294	143	17·413	0,08164	104
8	5·857	0,06401	71	12·527	0,05604	51
12	4·416	0,05326	47	10·266	0,04617	34
16	4·163	0,04790	35	11·371	0,04931	25
20	3·626	0,04346	27	8·944	0,03843	20
24	3·644	0,04497	23	12·232	0,04253	16
28	3·858	0,04144	19	10·080	0,03744	14
32	2·705	0,02823	17	5·625	0,02991	12
36	3·169	0,03248	15	7·414	0,02805	10
40	2·519	0,02668	13	4·450	0,02072	9
64	2·631	0,03829	8	5·595	0,02384	5
96	2·271	0,02199	5	2·029	0,01235	3

Size of Plots	Field III – Gen. Mean: 6,285 (332 trees)			Field IV – Gen. Mean: 4,410 (377 trees)		
	Variance raw data*	Variance transformed data	d.f.	Variance raw data	Variance transformed data	d.f.
—	23·423	0,18084	331	13·047	0,18211	376
—	13·275	0,06105	75	7·488	0,08686	86
—	8·682	0,04245	37	5·366	0,06219	42
—	8·778	0,04045	24	4·499	0,05439	28
—	7·646	0,03662	17	4·832	0,05296	20
—	7·520	0,03303	13	3·583	0,03216	16
—	7·558	0,03322	11	3·728	0,03828	13
—	6·727	0,02514	9	3·799	0,04151	11
—	7·780	0,02854	8	3·194	0,04189	9
—	6·323	0,02160	7	4·222	0,04918	8
—	7·176	0,02986	6	3·211	0,02655	7
—	5·143	0,02144	3	2·326	0,01536	4
—	7·042	0,02831	2	2·648	0,02405	2

Size of Plots	Field V – Gen. Mean: 6,590 (303 trees)			Total of fields: (2,008 trees) Gen. Mean raw data: 5,477 Gen. Mean transformed data: 3,57971			
	Variance raw data	*Variance transformed data*	*d.f.*	*Variance raw data*	*Variance transformed data*	*d.f.*	*c.v. (per cent)*
—	17·342	0,13661	302	19·156	0,17558	2·007	11,7
—	9·604	0,06014	72	11·539	0,08389	484	8,09
—	8·144	0,05010	35	8·465	0,05996	240	6,85
—	6·554	0,04463	23	7·183	0,05181	160	6,36
—	6·039	0,03483	17	6·975	0,04772	118	6,10
—	6·124	0,03325	13	6·194	0,03997	93	5,59
—	6·780	0,04191	11	6·789	0,04229	78	5,74
—	5.246	0,02867	9	6·292	0,03845	66	5,48
—	4·534	0,02903	8	4·826	0,03311	58	5,08
—	4·891	0,02976	7	4·039	0,03497	51	5,22
—	5·179	0,03723	6	3·714	0,02823	45	4,69
—	4·976	0,02588	3	3·996	0,02822	27	4,69
—	4·557	0,02325	2	3·227	0,02079	18	4,03

Data expressed in grammes of fresh beans.
* Variances of raw data are expressed in thousands of grammes.
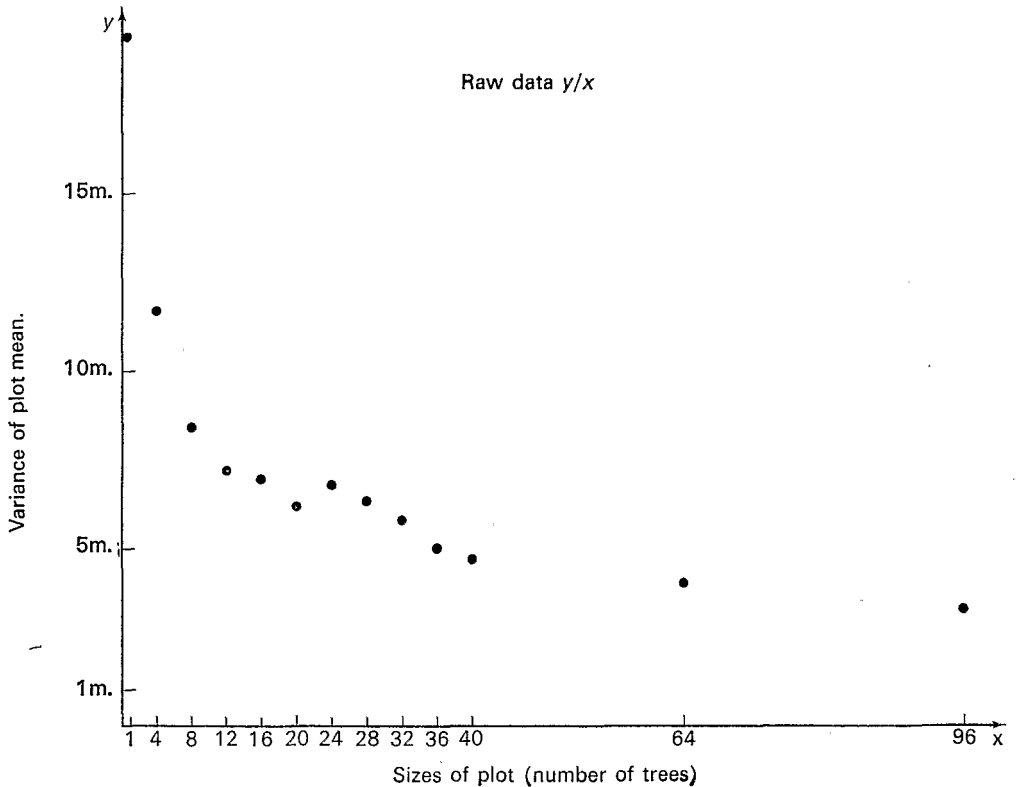d.f. = degrees of freedom.

Fig. 4   Relation between variance of raw data and plot size.

2*

Raw data

Examination of Table 2 shows that the variances for plots of the same size show a distinct variation from one field to another; the ratio of extreme variances ranges from 2·1 to 3·4, depending on the size of plot. They are significantly different.

Transformed data

The variances associated with plots of the same size show little variation from one field to another. If only those calculated with an adequate number of degrees of freedom are considered, the relative range varies, depending on the size of plots, from 1·33 to 1·65. They are no longer significantly different. This demonstrates once again, if indeed it is still necessary, the value of the logarithmic transformation.

Whatever the level of yield of the blocks or whatever the number of years used in the cumulation (after a certain minimum, the first two years of production show a variability such that it is not possible to analyse them reliably), the residual error variance on the plot mean (for a given size) becomes constant. This will allow us to study in a general manner the precision of an experiment of a given type.
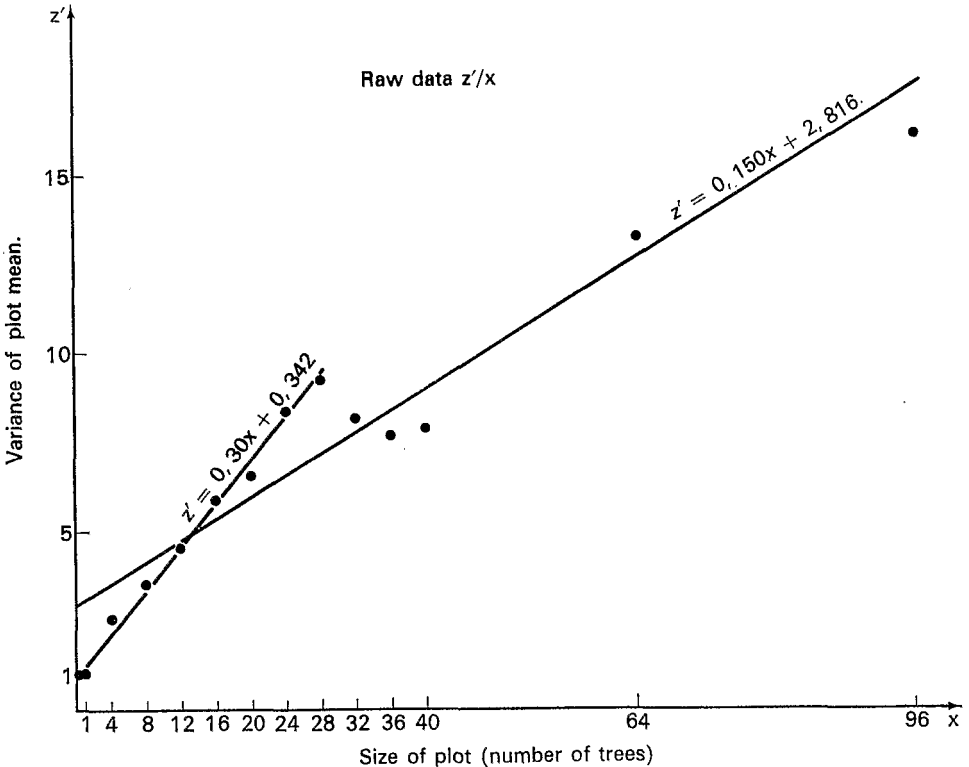

Fig. 5   Test of hyperbolic relationship between variance and plot size (raw data).

Study of changes in the variance of raw data in terms of plot size

Let us plot on the abscissa of Fig. 4, the sizes of plots (x) and on the ordinate the variance of the corresponding mean of the plots (y). The points plotted suggest a definite hyperbolic relation of the form: $y = a/x + b$.

To verify this, plot on the ordinate the variable: $z = y \cdot x$. If the relation suggested above is true, z must be a linear function of x: $z = a + bx$.

Calculating the regression will then allow us to define coefficients a and b. To simplify the calculations, we take as unit on the ordinate the value of the individual data: $19,156 \times 10^3$, thus obtaining Table 3.

Table 3

x	z	z' (*as new units*)
1	19,156,431	1
4	46,154,168	2·409
8	67,717,768	3·535
12	86,200,620	4·500
16	111,592,512	5·825
20	123,888,800	6·467
24	162,928,512	8·505
28	176,189,972	9·197
32	154,424,384	8·061
36	145,402,672	7·590
40	148,553,040	7·755
64	255,760,896	13·351
96	309,760,032	16·170

The points are practically in a straight line as shown for (z'/x) in Fig. 5. In point of fact:

correlation coefficient:	$r = 0.95$ (11 d.f.)
regression slope:	$b = 0.152$
regression line:	$z' = 0.152x + 2.816$
	$y = 0.152 + 2.816/x$

Fitting to a hyperbola is satisfactory and it is noteworthy that the same conclusion is reached here as for trees in traditional type plantations (No. 2): the variance of the plot mean does not tend towards 0 when the size increases indefinitely but towards a value corresponding to approximately 15 per cent of the individual data. If the distribution of trees within the plantations were strictly at random, i.e. if the hypothesis of independence of yield of neighbouring trees were correct, the variance inside those plantations would follow the general theoretical law: $y = a/x$ (a being the variance of the individual yields). The coefficient b measures the degree of association of trees with comparable yield.

As to the total yield of the trees, very similar values for the coefficient b ($b = 0.152$ and 0.16) were obtained from the two very different populations of cocoa trees, clearly showing the general nature of the phenomenon. It should also be noted that for pods only affected by black pod (*Phytophthora palmivora*)

the coefficient b takes a much higher value, which may be as high as 0·25, showing to some extent the existence of the phenomenon of contagion between a tree and its neighbours, but mainly the existence in the ambient environment of spots favouring the development of the parasite.

Referring back to the formula: $z' = 0·152x+2·816$, it will be seen that the fit is not satisfactory for the first few points. Examination of the graph shows that the curve passing nearest to the points is inflected towards the axis of the abscissa, i.e. it seems to tend towards an asymptote. It should also be noted that the last few variances, calculated with fewer degrees of freedom, are not very precise. The first 8 points are almost perfectly aligned:

$r = 0·998$ (6 d.f.)
$b = 0·30$
$z' = 0·30x+0·942$
$y = 0·30+0·942/x$

Fitting is now excellent. The higher value of b ($= 0·30$) so obtained points to the substantial influence of the relationship between trees with comparable yield extending up as far as plots of approximately 28 trees. Beyond this b decreases, indicating that the influence of the association of trees with similar
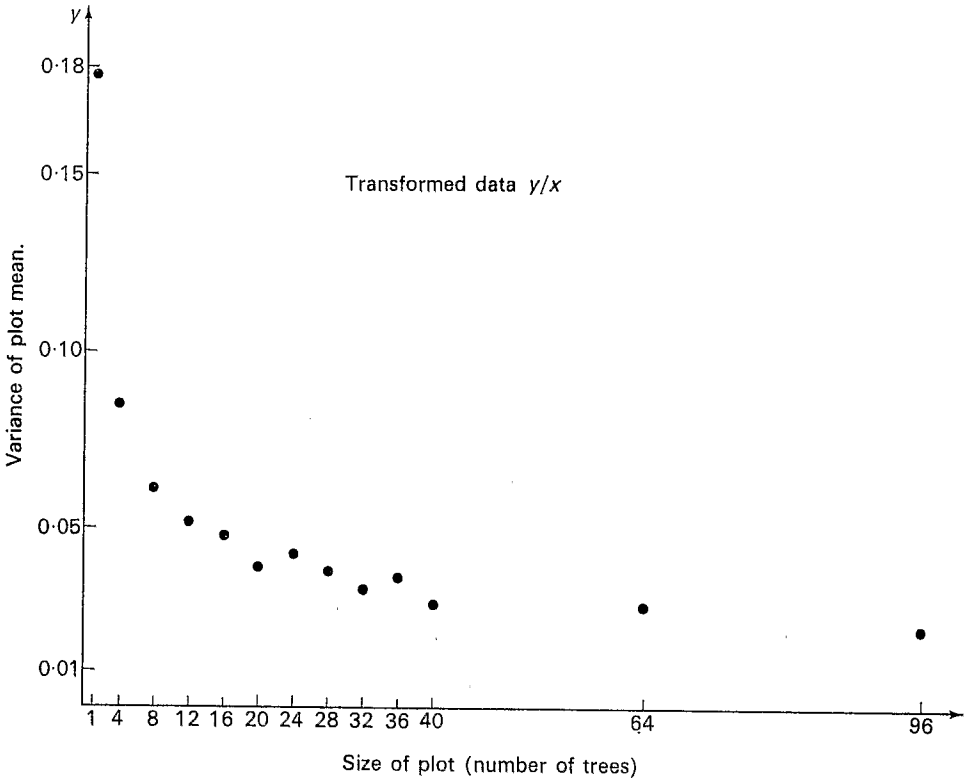

Fig. 6   Relation between variance of transformed data and plot size.

yield decreases as the size of the plots is further increased; this is logical: areas of identical fertility are small (heterogeneity of soil detail) and similarly the influence of an individual shade tree extends over a small area.

Study of pattern of the variances of transformed data in terms of plot size

Plot sizes (x) are plotted on the abscissa of Fig. 6, with the variance of the corresponding transformed plot mean (y) on the ordinate. Here again the corresponding points suggest a definite hyperbolic relation of the form: $y = a/x + b$. As before, let us examine the linear regression existing between $z = y \cdot x$ and x (see Fig. 7, z/x data transformed).

Table 4

x	z
1	0·17558
4	0·33556
8	0·47968
12	0·62172
16	0·76352
20	0·79940
24	1·01496
28	1·07674
32	1·05952
36	1·25892
40	1·12920
64	1·80608
96	1·99584

$r = 0·96$ (11 d.f.)
regression slope $= 0·01929$
$z = 0·01929x + 0·39748$
$y = 0·01929 + 0·39748/x$

Fitting to the hyperbola is good, and here too the variance does not tend towards 0 as x increases, but towards a value of 0·01929 which represents 10 per cent of the variance of the individual data. The degree of association seems to be lower after transformation; this is due to the fact that the points get closer together as the values get smaller, and to the relatively smaller importance given to the extreme values.

More clearly than with the raw data, a constant curvature of the best fitting curve is noted towards the axis of the abscissa. The influence of the degree of association of trees with comparable yield exerts a decreasing effect on the variance as the size of plot increases.

Fitting to a straight line is excellent only for the first 8 points:

$r = 0·99$ (6 d.f.)
regression slope $= 0·03276$
$z = 0·03276x + 0·19559$
$y = 0·03276 + 0·19559/x.$

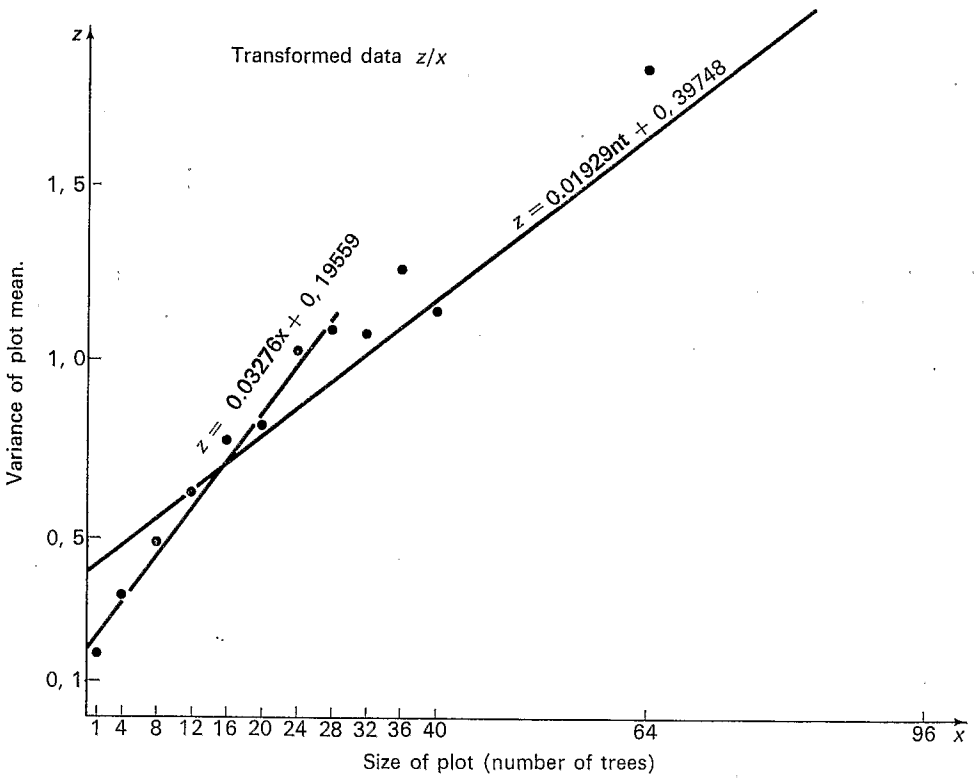The fitted values are then very close to the calculated values for these 8 points.

Fig. 7   Test of hyperbolic relationship between variance and plot size (transformed data).

Table 5

	Fitted value	Calculated value
$x = 1$	$y = 0.22835$	0.17558
$x = 4$	$y = 0.08166$	0.08389
$x = 8$	$y = 0.05721$	0.05996
$x = 12$	$y = 0.04906$	0.05181
$x = 16$	$y = 0.04499$	0.04772
$x = 20$	$y = 0.04254$	0.03997
$x = 24$	$y = 0.04091$	0.04229
$x = 28$	$y = 0.03975$	0.03845

In practice, *beyond 20 to 25 trees, an increase in the size of the elementary plot results in only a negligible decrease in the error variance, and there is absolutely no point in going beyond.*

Study of the precision of a given trial according to the design adopted

The precision of the general mean over all the fields under observation will be greater, the smaller the variance of this mean.

Suppose we have a group composed of N trees. We ask: what must be the elementary plot size in order to obtain the greatest precision of the general mean, given the above information?

Case where the trial does not require a border line around the elementary plots (trials with varieties, hybrids and clones)
In this case the fitted values are used for the variances.

Table 6

Elementary plot (number of trees)	Number of elementary plots in a field of N trees	Variance of the general mean
1	N	$\dfrac{0\cdot22835}{N}$
4	$N/4$	$\dfrac{0\cdot08166}{N/4} = \dfrac{0\cdot32664}{N}$
8	$N/8$	$\dfrac{0\cdot05721}{N/8} = \dfrac{0\cdot45768}{N}$
12	$N/12$	$\dfrac{0\cdot04906}{N/12} = \dfrac{0\cdot58872}{N}$
16	$N/16$	$\dfrac{0\cdot04499}{N/16} = \dfrac{0\cdot71984}{N}$
20	$N/20$	$\dfrac{0\cdot04254}{N/20} = \dfrac{0\cdot85080}{N}$
24	$N/24$	$\dfrac{0\cdot04091}{N/24} = \dfrac{0\cdot98184}{N}$
28	$N/28$	$\dfrac{0\cdot03975}{N/28} = \dfrac{1\cdot1130}{N}$
	etc. . . .	

Clearly, the complete randomization of all trees is superior to any other arrangement. As an example consider two designs set out in a field of N trees:

complete randomization;
complete randomized blocks with elementary plots of 20 trees (system adopted at Nkoemvone for the first comparative trials on hybrids).

Let d be the smallest significant difference between two consecutive means/hybrids after ranking in order of magnitude in the first case (total randomization), and d' in the second case (randomized blocks).

$$d = t \times \sqrt{2 \times \text{error variance of mean}} \quad \text{(Student } t\text{)}$$

$$d = t \times \sqrt{2 \times \frac{0\cdot22835}{N}}$$

$$d' = t \times \sqrt{2 \times \frac{0\cdot85080}{N}}$$

$$\frac{d'}{d} = \sqrt{\frac{0\cdot85080}{0\cdot22835}} = 1\cdot93.$$

Taking the calculated variances instead of the fitted ones, this gives: $d'/d = 2.13$.

In the second design, the smallest significant difference between transformed means must be approximately twice as much as the difference necessary in the first design. To what does this difference actually correspond?

Let M_1' and M_2' be two consecutive means after ranking in order of magnitude in the second design; and M_1 and M_2 in the first design. The foregoing ratio then becomes:

$$(\log M_2' - \log M_1')/(\log M_2 - \log M_1) \simeq 2$$

$$\log \frac{M_2'}{M_1'} \simeq 2 \log M_2 \simeq \log \left(\frac{M_2}{M_1}\right)^2$$

$$\frac{M_2'}{M_1'} \simeq \left(\frac{M_2}{M_1}\right)^2$$

(Note that the Ms are the geometric means obtained by simple re-transformation, and they show a skew with respect to the arithmetic means.)

If the significant ratio in the first system is 1.2, in the second system it will be 1.44. From 1.3 in the first system it will go up to 1.69. The differences are quite substantial and complete randomization is to be adopted whenever facilities permit—especially qualified personnel. Trials involving varieties, hybrids and clones should be set up using this design in all research stations.

If it is thought possible to control part of the heterogeneity of the experimental field by a block stratification, the complete randomization per block design should be adopted which will provide for:

a reduction of the random error;
obtaining continued advantage from complete randomization (minimum variance of the treatment mean).

Case where the trial requires a border line around the elementary plots (trials with fertilizers, density, shade etc.)

Table 7

Useful elementary plot (no. of trees)	Total elementary plot (no. of trees)	Number of plots in a field of N trees	Variance of the general mean
1	9	$N/9$	$2.055/N$
4	16	$N/16$	$1.307/N$
6	20	$N/20$	$1.307/N$
8	24	$N/24$	$1.373/N$
9	25	$N/25$	$1.362/N$
12	30	$N/30$	$1.472/N$
16	36	$N/36$	$1.620/N$
20	42	$N/42$	$1.787/N$
24	48	$N/48$	$1.964/N$
25	49	$N/49$	$1.989/N$
28	54	$N/54$	$2.147/N$

Theoretically the elementary plot of 6 trees gives the maximum precision per unit of experimental area. But other factors are taken into consideration in this type of trial.

Mass effect

In a fertilizer trial, for example, modification of the soil equilibria through application of fertilizers can be expressed only if it is carried out on an area of a certain size. In a shading experiment modification of the microclimate can be expressed only over a still larger area: here the effect of mass is very important.

In a fertilizer trial an elementary plot of 20 trees (4×5) for 6 useful ones (2×3) seems insufficient to create the required 'soil environment'. Furthermore, the ratio useful trees/total trees is unfavourable. This must be taken into account when the cost of the trial is a factor in the choice of design. An elementary plot of 30 trees (5×6) with 12 useful trees (3×4) seems to be the lower reasonable limit. The presence of various species of shade trees acting on small areas through physico-chemical modification of the soil, and establishment of a very localized microclimate, increase the degree of association of trees with comparable yield. In a trial started after complete deforestation, this degree of association certainly decreases, but in what proportion? Soil heterogeneity still exists, and is due to:

the microtopography;
the very variable depth of lateritic concretion;
the variable influence of forest tree species, which persists for several years after felling.

However, since the coefficient b decreases the optimum theoretical size of plot increases, and this is one more reason for adopting 12 trees as a lower limit. Furthermore, it seems that there is absolutely no advantage in going beyond 20 to 25 useful trees.

Here then we have a range of 12 to 25 useful trees within which to operate.

Availability of surface area, and availability of qualified personnel

If surface area not limited

If there is no problem of personnel, the work can be carried out in the upper part of the range and useful plots of 25 trees will be adopted. The number of replicates will depend on the precision desired.

If there is a problem of personnel, 25 trees can still be adopted but the number of replicates will depend on personnel and the precision of the trial will be reduced.

If surface area is limited

If there is no personnel problem, operations will be carried out in the lower part of the range and useful plots of 12 trees will be adopted giving the maximum precision per unit of experimental area (within the range). The number of replicates will be determined by the area available and a certain precision will result therefrom.

If in addition there is a problem of personnel, operations will be carried out in the middle of the range so as to:

occupy the available area;
adjust the number of replicates in relation to the available personnel.

Example: Precision of two designs

3^3 *design with elementary plots of 25 useful trees with 4 replicates.* The formula given above gives an approximate value for the variance of the plot mean (transformed data): 0·04058. There are 12 effective replicates for each main effect. Let M_1 and M_2 be means of results of two consecutive treatments after ranking in order of magnitude. They will be significantly different if:

$$\frac{\log M_2 - \log M_1}{\sqrt{2 \times \dfrac{0·04058}{12}}} > 2$$

$$\log \frac{M_2}{M_1} > 0·16446$$

$$\frac{M_2}{M_1} > 1·46.$$

The significant ratio is very high under the conditions adopted at Nkoemvone. For study of the interaction, the ratio of the means will have to be greater than 1·93! One can hope that after complete deforestation the variance of the plot mean will be distinctly reduced, and that consequently the smallest significant difference will also be reduced, but by how much is not known because data from use of such a cultivation technique are lacking.

3^3 *design with elementary plots of 12 useful trees with 8 replicates.* (Corresponding to a 25 per cent increase of experimental area over the preceding design.)

Taking the fitted variance of the plot mean: 0·04960, similar calculations lead to the following significant ratios:

for the main effect: $M_2/M_1 > 1·34$
for the interaction: $M_2/M_1 > 1·67$

Although lower than in the preceding case, these ratios are still relatively high.

This first point having been settled, it is now necessary to try to decrease the random error by considering the possibility of measuring the influence of some of the factors contributing to it.

Study of the correlation between yields of adjacent plots

In their study conducted with trees in traditional plantations, Marticou and Muller (1964) sought a variable satisfying the following requirements of statistical analysis:

normality of distributions;
identity of variances;
independence of observations.

They studied in particular the differences between the yields of plots during successive years and the differences between the actual yields of plots and the theoretical yields from these same plots, taking their intrinsic potential and the effect of the year's climate into account. This study showed that the climatic effects on yield, as well as the effects of an applied treatment are not quantities which combine additatively with the potential yield of the tree but that these effects, which are more important on high-yielding trees, might be a multiplicative factor of yield.

This led them to suggest, in an experimental scheme in which a control plot is associated with each treated plot, the following particular model:

$P_1 = P_0 k_1 \varepsilon = P_0(T_1/T_0)I\varepsilon$, where
P_0 = yield without any treatment in year 0 of the plot to be treated in year 1;
P_1 = yield of the same treated plot, year 1;
T_0 = yield of the control plot in year 0;
T_1 = yield of the control plot in year 1;
I = a multiplicative factor characterizing the effect of the treatment;
ε = a random normal variable with a mean of 1.

Converting to logarithms

$$\log P_1 - \log P_0 = \log T_1 - \log T_0 + \log I + \log \varepsilon$$

(log ε assumed to be a random normal variable with zero mean).

This amounts to determining whether between year 1 and year 0, the increase in yield of the treated plot is greater than the increase in yield of the adjacent control plot, the data being expressed in logarithms. This is actually a simplified covariance model for which it is assumed that the slope of the regression line is equal to 1.

A study to test this hypothesis was made (Lotodé, 1971) using yields of trees forming part of the Nkoemvone Station collections. The study was limited to plots of 12 trees.

Seventy-nine independent pairs of adjacent plots, each comprising 12 trees, were set up (pairs 1–2, 3–4, etc.). The study of the correlation between yield means (transformed into logarithms) of adjacent plots led to the following results (note that the transformation of the data normalizes the distribution and makes this study possible):

$r = 0.344$ for 77 degrees of freedom, which is very highly significant (r at threshold 0·01 and for 77 d.f. $\simeq 0.29$);
regression slope, $b = 0.392$.

The same study using the raw data shows a stronger relationship ($r = 0.53$, $b = 0.57$). But in this case the conditions required for a valid analysis were not present. The distribution of raw means for the plots (although calculated on the basis of 12 observations) is quite distinctly skew and non-normal.

Let x be the data of the set formed by an element of each pair and y the data of the complementary set. With the data transformed the following is obtained:

$$s_x^2 \simeq s_y^2.$$

Correlation coefficient $r = s_{xy}/s_x s_y$.

Since $s_x \simeq s_y$, $r \simeq s_{xy}/s_x^2$. This latter expression represents the regression slope. We therefore have $r \simeq b$. This is indeed found in our case: 0·344 and 0·392.

Let us compare the following two designs set out on the same area, the elementary plots having the same size (12 trees):

supplementing each treated plot with an adjacent control plot with a view to subsequently undertaking an analysis of covariance (each treatment is applied on n plots);

conventional design, comprising therefore a double number of replicates ($2n$).

1st case

The correlation coefficient was determined between means of adjacent plots, $r = s_{xy}/s_x s_y = 0·344$ (s_{xy} = covariance $x/y - s_x$, and s_y = standard deviations of x and y).

In the analysis of covariance the residual sum of squares of the deviations is decreased by the quantity:

$$Q = \frac{(S_{xy} \text{ residual})^2}{S_x^2 \text{ residual}}$$

(where S represents a sum of squares of deviations or of products of deviations).

$$Q = r^2 \cdot S_y^2 \text{ residual}.$$

The sum of squares of the residual deviations becomes:

$$S_y^2 \text{ residual} - r^2 \cdot S_y^2 \text{ residual} = (1 - r^2) S_y^2 \text{ residual};$$

and the residual mean square becomes:

$$(1 - r^2) s_y^2 \text{ residual}.$$

The variance of the mean will be approximately:

$$\frac{(1 - r^2) s_y^2 \text{ residual}}{n}$$

(a correction should in fact be made to take into account the additional inaccuracy due to the use of the regression coefficient b, but it can be disregarded in this approximate calculation).

2nd case

Here the residual mean will be calculated with a greater number of degrees of freedom, but it must be only slightly different from the foregoing residual s_y^2.

Since the two designs are assumed to have been set up on two identical fields, the random error variance will be constant whatever the design may be.

The variance of the mean will then be:

$$\frac{s_y^2 \text{ residual}}{2n}.$$

The relative efficiency of the 2nd design compared with the 1st will correspond to the ratio of the variances of the means:

$$\frac{s_y^2 \text{ residual } (1-r^2)}{n} \bigg/ \frac{s_y^2 \text{ residual}}{2n} = 2(1-r^2) = 2-2r^2.$$

This shows that the 2nd design will be more advantageous if $r < \sqrt{2}/2$, or $r < 0.7$, and vice-versa.

Therefore, whilst there is a positive correlation between the annual yields of adjacent plots, the analysis just made shows that the correlation coefficient is not large enough (only about 0·35) for the model proposed to be applicable in practice. In our case the relative efficiency, calculated without taking into account the numbers of degrees of freedom, is 1·76.

Admittedly, it may be supposed that by decreasing the plot size and by taking as a variable not the total yield of the trees, but an element more closely linked with the microclimate such as the percentage of pods affected by black pod, the correlation coefficient might be increased and the possibility of applying the model suggested by Marticou and Muller might come closer. But it seems improbable that the coefficient r would reach the high values required to allow this model to be used.

Problems of missing individuals due to accidental losses

A missing individual in an elementary plot leads to a *perturbation* in the plot data which must be measured and introduced into the analysis if the mortality cannot be a function of the treatment (in the opposite case, the yield adjusted to the initial total number of trees will be used in the analysis). The best way to proceed is to use the following analysis of covariance. Operations are carried out on the mean of plots calculated per surviving plant and this data is adjusted by the number of missing individuals per plot. The technique can be improved by weighting each missing individual by the time elapsed since the death of the tree; for example, by attributing a value of 5 to a tree missing for 5 years, a value of 2 for a tree missing for 2 years, etc. The covariance makes it possible to calculate over the entire trial the average influence of one missing individual, of two missing individuals, etc., on the yield per plot in relation to the number of surviving trees and apply the best possible correction. (It should be noted that the influence of the missing individuals will vary according to whether they are arranged in a group or dispersed in the plot. This cannot be taken into account.)

Pearce (1953, p. 74 *et seq.*) considers this procedure to be the best.

The problem of missing trees does not arise in the same way with the completely randomized design. The influence of a missing individual will be distributed over several hybrids, 8 at best. The mean for a treatment will therefore be less perturbed than in the block design. However, the problem can be solved in a similar manner. A pseudo-variate will be used, as follows.

For each tree the number of adjacent missing individuals (i), and for each of those, the number of years elapsed since the death of the tree (ai) are noted. The independent variable connected to the yield of the tree is Σai. The best possible relationship could moreover be sought between the two variables by attempting to find an expression which could be more representative of the competitive effect than Σai. This expression will be that which minimizes the random error in the analysis of covariance. This analysis will make it possible to calculate the average influence of 1, 2, 3, etc., adjacent missing individuals on the yield of a given tree. Whether missing individuals are arranged in groups or dispersed is no longer important here.

Conclusions on Part II

In conclusion, the positive results obtained in Part II can be summarized as follows.

Size of useful elementary plot

Border lines not necessary. Total randomization must be adopted tree by tree. If control of the heterogeneity of the experimental field is possible by block stratification, randomization can be done within each block.

Border lines necessary. A range of 12–25 trees is recommended, as follows:

surface area not limited:
    personnel and funds sufficient . . . 25 trees
    personnel and funds limited . . . 25 trees (however, if the number of
    replicates is limited this factor will determine the precision of the trial.)
surface area limited:
    personnel and funds sufficient . . . 12 trees
    personnel and funds limited: operation will be carried out at the midpoint
    of the range, so as to occupy the available area and adjust the number of
    replicates to the available personnel.

Design of the experiment

The conventional design will be adopted but the random error can be definitely decreased by introducing into the analysis the following factors whose effect can be measured:

number and date of observation of missing individuals per elementary plot, in all cases;
measurement of the trunk at 20 cm above the collar, carried out one and two years after planting, at which time this variable is not influenced by the applied treatments (as will be seen in the following study).

Part III. Search for a criterion other than yield per tree: relation between measurement at a given age and subsequent cumulative yield .

Working with a perennial plant, whose productivity becomes manifest only after several years, comparative studies between cocoa varieties are very lengthy. Attempts have been made to find a criterion, other than the actual yields of the trees, which would estimate productivity by representing these yields, and which by appearing during the very first few years after planting would make it possible to form an opinion about the varieties under study (clones or hybrid families) within a much shorter time.

In comparative experiments with hybrids started in 1964 at the Nkoemvone Station and comprising 40 families divided into elementary plots of 20 trees with 5 replicates (a total of 4,000 trees), measurement of the diameter of the trunk at 20 cm above the collar were made at 11 and 23 months after planting.

Two analyses of covariance were attempted, taking as dependent variable the mean accumulated yield per elementary plot, calculated per tree surviving 5 years after planting (transformed into logarithms) and, as independent variable, the mean per plot of the trunk measurements at 11 months (a) and at 23 months (b).

(a) In the first analysis (Table 8), the regression slope ($b = 0.69$) is significantly different from 0 ($F = 17.3$). The linear correlation coefficient is 0.56. The residual mean square is reduced from 0.05464 to 0.037795, which corresponds to a reduction of 31 per cent of its value.

(b) In the second analysis the regression slope ($b = 0.43$) is significantly different from 0 ($F = 38$). The linear correlation coefficient is 0.71. The residual mean square is reduced from 0.05464 to 0.02749, which corresponds to 50 per cent of its value.

The relationship is more interesting at 23 months. It seems therefore that this is the independent variable that should be used in the analyses. (An analysis carried out with an increase of the diameter between 11 and 23 months gave results of the same order of magnitude.) The calculations repeated six years after planting led to similar conclusions. It seems however that the increase in diameter of the trunk between 11 and 23 months might be a more interesting dependent variable than the diameter at 23 months. Whatever the case may be, both analyses should be performed so as to take into consideration the one leading to the greatest reduction of the residual mean square.

The analysis of covariance can be carried out only if the auxiliary variable is not influenced by the treatments. In this case, however, the 'mean diameter' variable is influenced by:

the mean fertility per plot;
the factor under study: the hybrid.

In fact the analysis of variance of the mean diameters shows that the differences are significant between hybrids. In this case, to reduce all the diameter/hybrid

Table 8.

First analysis of covariance
(diameters at 11 months)

Source of variation	Degree of freedom	Sx^2	Sxy	S_y^2	$\dfrac{(Sxy)^2}{Sx^2}$	S_y^2 reduced	D.F.	M.S.
Family	39	4·9596	2·58310	9·92490	—	8·65922	39	—
Block	4	0·4347	0·61660	3·08800	—	2·42395	4	—
Random error	156	5·5108	3·83294	8·52386	2·66593	5·8793	155	0·03779
Family+random error	195	10·4704	6·41604	18·44876	3·93161	14·51715	—	—
Blocks+random error	160	5·9455	4·44954	11·61186	3·32998	8·28188	—	—

Second analysis of covariance
(diameters at 23 months)

Source of variation	Degree of freedom	Sx^2	Sxy	S_y^2	$\dfrac{(Sxy)^2}{Sx^2}$	S_y^2 reduced	D.F.	M.S.
Family	39	14·0613	6·91459	9·92490	—	6·55770	39	0·16815
Block	4	5·8366	3·97441	3·08800	—	0·67430	4	0·16857
Random error	156	23·0308	9·90897	8·52386	4·26332	4·26054	155	0·02749
Family+random error	195	37·0921	16·82356	18·44876	7·63052	10·8124	194	—
Blocks+random error	160	28·8674	13·88338	11·61186	6·67702	4·93484	159	—

means to the general mean would result in the intrinsic value of the hybrid, which precisely shows variable growth in its early stages, not being taken into account.

In trials in which the elementary plots have the same genetic constitution, the mean diameter per plot will be a reflection of the mean fertility per plot only. The analysis of covariance indicated will then be very beneficial and it should be carried out whenever possible. In a mineral fertilization trial beginning when the trees start to yield, measurements made before the first treatment at 11 and 23 months, for example, could be used as indicated. This is indispensable if reasonable significant differences are to be obtained.

The significant ratios will in fact be decreased, very approximately, to the following values (in specially prepared forest, as previously):

1st design: $(3^3) \times 4$ (elementary plot of 25 trees)

$\qquad$ main effect $\qquad$: $\qquad$ 1·31

$\qquad$ interaction $\qquad$: $\qquad$ 1·58

2nd design: $(3^3) \times 8$ (elementary plot of 12 trees)

$\qquad$ main effect $\qquad$: $\qquad$ 1·23

$\qquad$ interaction $\qquad$: $\qquad$ 1·44

In the comparative trials with hybrids, families and varieties, the relationship indicated could allow us to make a first rapid choice by analysing the measurements without waiting for the plots to start producing.

It is not unreasonable to think that other criteria for estimating the effects of applied treatments will be found providing more rapid knowledge of the phenomena under investigation.

Part IV. Case of experiments with fungicides against *Phytophthora palmivora*: study of the intrinsic fungicidal value of a formulation

The foregoing study on the use of actual yields of trees as a criterion for estimating the effects of an applied treatment, or the value of different varieties of trees, was carried out on cumulated yields over several years. This presupposes a long-term experiment, which is usual when it involves the study of cultivation techniques (trials with fertilizers, spacing, cover or shade plants, planting procedures, etc.) or trials on different varieties, the results from which cannot reasonably be expected after only one season. When the value of hygienic measures, especially anticryptogamic treatments, integrated with disease-prevention measures forming current cultivation practice are to be studied, it is also usual to form an opinion on their effects on yield at the end of the season in spite of the difficulties that this involves. Some research workers have considered using covariance on yields in years preceding the treatments: one year is

insufficient because a tendency towards alternation of high and low yield; two years may be sufficient, but then the length of time required becomes prohibitive for this type of study.

When only the intrinsic fungicidal value of a formulation is to be estimated, the criterion of annual yield is even less readily usable: the phenomenon of contagion which acts differently on the pods (whose number and distribution can vary greatly from tree to tree) increases the experimental error, so that only large differences appear as significant. All trials made up to now have shown this to be true.

Now the intrinsic fungicidal action of a formulation against *P. palmivora* does not modify the productivity of the trees, but shows an effect only on the already existing yield. Instead of estimating its value by comparing yields of treated and non-treated trees, it can be estimated by taking as a criterion the percentages of diseased pods as compared with the numbers of treated and untreated pods.

But since a percentage has an intrinsic significance varying with the number on which it is calculated, and since the phenomenon of contagion acts differently as noted above, depending on the degree of aggregation of the pods, how then in an elementary plot can a weighted mean percentage be calculated, knowing that each tree in the group carries different numbers of pods with different distributions? This difficulty is practically insuperable and means that one has to forgo using the elementary plot grouping of several trees.

To facilitate the analysis without introducing too difficult mathematics – which still remains to be developed – we have been led to regard the elementary plot no longer as a group of trees but as a group of pods carried by a single tree, which is equivalent to stretching to an extreme limit the principle of reduction of plot size arrived at above (Muller, Lotodé and Njomou, 1969). This miniaturization of the experiment allows us to eliminate or reduce the effects of a certain number of heterogeneity factors.

Choosing plantations which appear favourable for the development of black pod so as to operate in the best conditions for contamination and development of the disease, it is possible to adopt a design of paired companions in such a way that:

the heterogeneity of the environment is largely eliminated, both trees of the pair being close to each other in the same environment (particularly having the same amount of shade);

the genetic heterogeneity is reduced, both trees of the pair carrying pods of the same type (same general shape, same colour);

the phenomenon of contagion is identical, both trees in the pair carrying the same number of pods, of the same size, and distributed in the same way over the same length of trunk (if necessary, certain pods can be eliminated in order to achieve these conditions). Only sufficiently large pods, which are not too near maturity, are retained so that they will not disappear during the experimental period through physiological wilt or maturation.

In order to minimize differences which may exist among natural sources of contamination – principally the soil surrounding the trees – five pods affected by black pod in the process of development and covered with *P. palmivora* fructifications are placed in a ring approximately 40 cm from the foot of each tree. An abundant and homogeneous contamination is thus provided. Only this source of inoculum must exist: when starting the experiment pairs of trees must be chosen without infected pods, because these could already have contaminated pods which apparently seem healthy. The trees close to the pair are subjected to severe plant hygiene.

Within each pair, the tree receiving the test fungicide is selected at random, the other receives the reference fungicide.

Observations are made every 15 days. Healthy and diseased pods per tree are counted. At each observation:

affected pods and cherelles appearing on the trunk are eliminated;
fungicides are applied;
the inoculum source is renewed.

The efficiency of a given treatment can be estimated by comparing the percentages of diseased pods obtained from the two treatments.

Under the optimum conditions described, percentages calculated on the same number of pods for the trees in all the pairs can be analysed after angular transformation homogenizing the variances (arc sin $\sqrt{p/100}$).

The technique of analysis will be that of Student's paired comparisons, which is the conventional parametric test. If xi is the difference observed between the transformed data within the ith pair, the estimate of the variance of this difference will be:

$$S_e^2 = \frac{\sum i(xi - \bar{x})^2}{k-1}$$

($\bar{x}$ being the average of the differences for all the pairs, and k being the number of pairs).

The two treatments are considered as significantly different if $\bar{x}$ is different from 0, i.e. if the value found for $t = \dfrac{\bar{x}}{s_e/\sqrt{k}}$ is greater than the value given in Student-Fisher's table for $k-1$ degrees of freedom.

The study of the precision of the design was made by analysing all the data collected from 17 experiments in which 10 pods were kept per tree for all the pairs.

In most cases, the action of a test fungicide was compared with copper oxychloride. If xi is the difference between the transformed percentages of the ith pair, and if k pairs are studied in a given experiment, the residual mean square representing the experimental error will be represented by:

$$s_e^2 = \frac{\sum i(xi - \bar{x})^2}{k-1}$$

The percentages employed were those observed on the 45th day after the beginning of the experiment.

Since the 17 residual mean squares obtained (k varying from 12 to 20 depending on the experiment) constituted a homogeneous whole, it was possible to calculate a residual mean square comprising a large number of degrees of freedom (282) which might therefore be used in calculating the precision: $s_e^2 e = 800$.

It was seen that the mean difference $\bar{x}$ observed between two treatments is significant when:

$$t = \frac{\bar{x}}{s_e/\sqrt{k}} > 2$$

or

$$\bar{x} \cdot \sqrt{k} > 2s_e.$$

Estimating se by $\sqrt{800} = 28$, the approximate relation is obtained: $\bar{x} \cdot \sqrt{k} \simeq 56$.

$$k = 3 \cdot 136/\bar{x}^2.$$

A smaller significant difference is therefore chosen simply in order to obtain the approximate number of pairs to be adopted. For example, if $\bar{x} = 10$ is chosen, 32 pairs must be adopted, for $\bar{x} = 15$, $k = 14$, etc.

In practice, to what does a smaller significant difference of 10 in the transformed data means correspond? It can be seen that after re-transformation that it corresponds approximately to a mean difference of:

> 10 per cent in the 20 per cent zone
> 15 per cent in the 35 per cent zone
> 17 per cent (maximum) in the 50 per cent zone.

This seems reasonable and it can be assumed that approximately 30 pairs are sufficient in practice to differentiate between fungicides (the boundary value t is effectively close to 2 in this case and no readjustment is necessary).

If a fixed number of pods per tree is adopted, the locating of pairs in the plantation will be very arduous even when some of the fruits are eliminated so as to provide, for example, a total of 10. It has been seen that this condition was imposed by the use of the analysis of variance described.

There exists a non-parametric analysis technique which allows one to disregard this constraint. If the number of pods does not vary much, for example if it is close to 15, contagion will not be substantially different from tree to tree. The data can then be analysed according to the method of Cochran (1954).

Cochran's test

For the ith pair, let:

$ni1$ and $ni2$ be the number of pods retained on the 2 trees;
$pi1$ and $pi2$ the percentage losses;

$\widehat{pi}$ the mean calculated percentage for the 2 trees $(\widehat{qi} = 1-\widehat{pi})$;

$di = pi1-pi2$, the differences between the two percentages;

$$wi = \frac{ni1 \times ni2}{ni1+ni2} \quad \text{and} \quad w = \sum iwidi \,.$$

The mean weighted difference is calculated for the pairs as a whole:

$$\bar{d} = \frac{\sum iwidi}{w};$$

Cochran shows that the standard deviation for this variable is:

$$s\bar{d} = \frac{\sqrt{\sum iwipiqi}}{w}$$

$$Z = \frac{\bar{d}}{s\bar{d}} = \frac{\sum iwidi}{\sqrt{\sum iwipiqi}} \quad \text{is then simply calculated,}$$

which follows a standardized normal distribution if the percentages are not different.

An estimate of the precision cannot be calculated using this non-parametric method. Cochran's method is less precise than Student's paired comparison method, but since the number of pods varies around 15 the percentages will have a greater intrinsic value and it can be assumed that the precision will be of the same order as that of Student's pairs with constant sample numbers equal to 10. The important point is that the locating of pairs will be greatly facilitated by adopting this technique of analysis.

Part V. Comparative study of cocoa clones as regards their susceptibility to *Phytophthora palmivora*

For the reasons given previously (number and arrangement of pods highly variable from one cutting to another, microclimatic heterogeneity of the environment), it is very difficult to see how to set up an experimental design in the field with the object of comparing the susceptibility of cocoa clones to *P. palmivora* (data from which reflecting natural infection) could be correctly analysed. One is led therefore naturally to adopt a technique involving experimental infections.

Working in the laboratory with detached pods under perfectly controlled environmental conditions and with the same strain of parasite provides the best solution to the statistical problem involved. However, certain observations weigh in favour of doing this work in the field in a natural environment with living pods still attached to their trees and therefore in a state of total physiological integrity, thus ensuring that exact clonal reactions have taken place with respect to the parasite.

The main objection to this field technique is that environmental factors are not controlled during the experiment.

Producing a large number of infections for a given clone, and repeating these infections during a season and for several years, should eliminate possible effects of the external environment. In addition, using on each infection occasion a fresh strain of the parasite taken from the site where the test is being made, will ensure that tests are being made with an inoculum representing all the strains which can exist locally and which determine the local degree of severity of the disease.

Infection is brought about by depositing a standard drop of a calibrated zoospore suspension. This technique has been used in Cameroon for several years (Blaha, unpublished, and see Chapter 21), and a considerable amount of numerical data has been collected. The following are recorded for each pod:

success or failure of infection;
time of appearance of diffuse spots, mottled, veined brown marks, fructifications of the parasite;
size of the spots each day.

The first of these data enable one to calculate the percentage of successful infections representing the vulnerability of the pod cortex to fungus attack.

The last set of data allows us to estimate the area of the spot on a daily basis and to calculate the mean daily progress, which has been the subject of a comparative study.

The mean progress between the 4th and 10th day after infection was chosen as it had the smallest percentage standard deviation. These data for a single clone are highly dispersed. The intraclonal variance is large, and it is found to increase with the mean so that the standard deviation of these data is a linear function of the means. Consequently, the logarithmic transformation was used. Analysis shows that the precision is acceptable with at least 80 pods per clone.

Part VI. General conclusion

Finally we emphasize once again the great complexity of experimentation with cocoa in the field, whether for studies of an agricultural nature or more specialized studies for improving methods of controlling *P. palmivora*, the screening of efficient fungicides, or experiments to determine the susceptibility of cultivars with respect to the parasite.

We have described our own experiments carried out in Cameroon. We do not in any way consider the conclusions to which they have led as panaceas but as approaches designed to overcome the difficulties which the experimenter faces, and we hope that they may serve as a basis or starting point for other methodological research, which will in turn contribute to discovering the best solution to this delicate problem.

Aberdeen, 1973

3

Problems of Experimentation with Cocoa Trees

By R. LOTODÉ and RAOUL A. MULLER

B3702

B3702
83/72/3762