

STATISTICAL GUIDELINES FOR COMPARING COMMERCIAL OIL PALM VARIETIES

BREURE, C J*

ABSTRACT

Samples of 42 *dura* x *pisifera* progenies of Ekona, AVROS, Ghana and Nigeria oil palm varieties, each involving six *pisifera*, were used to determine the minimum sample size of their parents for obtaining statistical yield differences between varieties. The 95% confidence interval of differences between varieties decreased sharply and, hence, the precision of the test improved, when increasing the number of progenies from 6 to 30 and *pisifera* (involved) from 1 to 4 and then levelled off to a steady decrement. As a test of differences the confidence interval has only a power of 50%, being the probability that a false null hypothesis of zero difference between varieties is rejected. With a power of 80%, a sample of 60 progenies in three replications detects, depending on the variance of the *pisifera* parents, a significant yield difference of 4.5% between Nigeria and AVROS and 7% between Ghana and Ekona with 6 and 9 *pisifera* in the sample, respectively. The article argues to incorporate optimal planting density as a crucial trait in selecting oil palm varieties. This study offers a method to obtain unbiased progeny samples, along with a field lay out for bulk recording of progeny plots.

Keywords: plant breeding management, oil palm variety test, power of the test, size of progeny sample.

Date received: 19 April 2016; **Sent for revision:** 20 April 2016; **Received in final form:** 20 January 2017; **Accepted:** 21 January 2017.

INTRODUCTION

Selection of Oil Palm Varieties

A key management decision is to select oil palm planting material with the highest yield for a certain environment. Planting material is presently all of the *tenera* fruit form, which is produced by pollinating *dura* female parents by *pisifera* male parents. Whereas seed producers use thousands of *dura* female parents, usually of the Deli type, only a restricted number of *pisifera* male parents are involved in producing *tenera* seed. ASD (in Costa Rica), which is a well-known supplier of oil palm seed, uses *pisifera* parents from various

distinct origins. Planting material derived from a certain *pisifera* origin, crossed with the same *dura* population, is here referred to as a variety.

The choice of a variety may be biased by the price of the seed. Furthermore, superficial comparisons of yield are sometimes made from large blocks with planting material from various seed suppliers. These blocks, which are sometimes planted at time-different periods, may consist of a mixture of *dura* x *pisifera* progenies or even a mixture of varieties provided by one supplier. Even when, for example, government organisations compare samples of varieties from several seed agencies, the outcomes of the test are likely exaggerated. Some suppliers may offer only the best crosses. As a consequence, the samples of progenies do not estimate the true (mean) yield of the population. Another shortcoming is that the number of progenies tested is too small for statistical backing of the results (Corley and Tinker, 2016; Isa *et al.*, 2009). A final reservation may be that the obtained results exclusively apply to a certain

* PT ASD-Bakrie Oil Palm Seed Indonesia, Jl Besar Air Joman, Pasar IX, Kisaran 21263, North Sumatra, Indonesia. E-mail: breurekees@gmail.com

environment. It is therefore timely for a statistically correct procedure to select the best oil palm varieties, as shown in the present study.

Selection Traits

The main interest focuses on the oil yield per hectare, which is the joint result of: (1) bunch yield per palm, (2) the number of palms per hectare (planting density) and (3) the oil extraction rate of the bunches. The genetic potential of oil extraction is unaffected by environmental conditions (Breure, 2003). Moreover, seed suppliers usually provide this information on request. Optimal planting density, which is the density giving the highest cumulative yield per hectare, can be derived from the leaf area index (LAI), *i.e.* the total leaf surface per unit land area. In oil palm, LAI is calculated from the number of palms per hectare, the number of leaves per palm and their mean leaf area. The LAI that provides the highest yield per hectare is specific for environmental conditions (von Uexkull *et al.*, 2003) and for practical purposes an optimum range of 5.5 to 6 is used (Breure, 2010). The LAI value at mature crown size, as reached about 10 years after planting, would give the most realistic estimate of optimal planting density. As in commercial plantings the number of leaves per palm is known (Gerritsma and Soebagy, 1999), the optimal planting density is only derived from a reliable estimate of the final size of mean area of the leaves. Traits that are needed to test the varieties of the present study are, therefore, restricted to the bunch yield per palm and leaf area at maximum crown size.

Random Sampling of Progenies per Variety

Testing all thousands of possible crosses between *dura* and *pisifera* palms (progenies), being the entire population of a variety, is unrealistic for some reasons. In the first place, it is not economical. Moreover, it is too laborious to obtain all the data of specific traits that are of relevance. Our strategy is, therefore, to cut back on the size of the samples. At the same time, it should be kept in mind that such a reduction does not affect the statistically reliable test procedure. Samples of progenies are randomly drawn, which estimate the population of, for example, the mean yield of a certain variety. Besides the total number of progenies, the number of *pisifera* involved in the crossing is crucial (*pisifera* of a certain origin determine the specific characteristics of varieties, whereas the set of *dura* female parents is the same). As a consequence, one has firstly to take a random sample of *pisifera* parents and secondly a random sample of *dura* parents crossed per *pisifera*. No publications on oil palm are known to the author that specify the size of the random samples of crosses per variety.

Database of Our Study

As database of our study, we used records from *dura* × *pisifera* progenies involving AVROS, Ekona, Ghana and Nigeria *pisifera* origins from ASD (Costa Rica), crossed with the same Deli *dura* population. Per origin a random sample was taken of six *pisifera* and, further, a random sample of seven *dura* per *pisifera*, resulting in a total of 42 progenies per sample. Each progeny was planted at three replicates and recorded for yield and leaf area at maximum crown size. From these data, we calculated per variety the variance components between *pisifera*, the variance components between *dura* within a *pisifera* and, finally, the variance components of error between replicates of progeny plots.

Confidence Intervals of Differences in Selection Traits between Varieties

The difference between the sample means, for example of the yield, obtained from two varieties is an estimate of the difference between the population means. Along with the standard errors, one can test whether the difference between varieties is zero at a level of significance of, for example, 5%. Equivalent to testing (with a significant level of 5%) differences between varieties one can calculate the 95% confidence interval of differences between population means. If the value zero does not fall in the 95% confidence interval, the null hypothesis must be rejected. Instead, the alternative hypothesis is accepted, implying that the means of the two varieties are unequal at a significance level of 5%.

Sample Size

It should be clear that with increasing the number of observations (here the number of progenies under test) the length of the confidence interval decreases, and, hence, accuracy of the test improves. As mentioned before, for practical purposes our interest is to limit the number of progenies under test. For this reason, we studied the effect of sample size, in terms of the number of progenies and number of *pisifera* parents involved in the sample, on the length of the confidence interval when testing of differences in yield between varieties.

Power of the Statistical Test

As the 95% confidence interval is equivalent to a test of zero difference at a significance level of 5%, it has a power of only 50%. When increasing the power of the test to 80% or even more, *i.e.* the probability that the test correctly rejects a false null hypothesis of equality between two varieties, the size of sample of progenies should be enlarged. This increment

could be studied by using the yield records of the database. Besides the number of *pisifera* in the sample, the number of progenies depends on the number of replicates in order to increase the power of the test.

Practical Range of Number of Progenies for Comparing Varieties

To compare the bunch yield of various varieties, Breure and Verdooren (1995) proposed planting samples of progenies from each variety in plots of two half-rows separated by an inspection path in the middle of a plantation unit. With standard distances of 1000 m between sub-main roads and 333 m between roads, at a density of 135 palms ha⁻¹, a total of 4500 palms can be planted in one unit (Turner and Gillbanks, 2003). So, one unit can accommodate a maximum of 120 progeny plots of 36 palms. Practical options are therefore to test samples consisting of 15; 30; 60 or 120 progenies per variety.

Objectives

The article aims to improve some conventional procedures of comparing oil palm varieties. First, we studied the effect of the number of progenies and the number of *pisifera* involved on the length of the 95% confidence interval and, hence, on the accuracy of the test. Since this was calculated with a power of 50% only, the length of confidence intervals is not sufficiently reliable to determine significant differences between varieties. We therefore calculated, with a larger power of minimal 80%, the minimum number of *dura* and *pisifera* parents of the progenies to obtain a certain significant yield difference between pairs of varieties. Novel in the present study is that, for each set of parents with two or three replicates, the power of testing the differences between paired varieties was included.

MATERIALS AND METHODS

Plant Material

The present article compares *dura* × *pisifera* progenies derived from four *pisifera* origins (varieties), involving a total of 39 *pisifera* parents: AVROS (15); Ekona (10); Ghana (8); and Nigeria (6). Each *pisifera* male parent pollinated about nine *dura* female parents. From each variety, a random sample was taken of six *pisifera* and then within each *pisifera*, a random sample of seven *dura*, resulting in a total of 42 progenies per sample. Random sampling was done with the Excel function *Random Permutation*. The progenies were arranged in a split-plot experiment with varieties as main plots and progenies as sub plots in three replications and 16 palms per plot (Breure, 2006). The progenies were

planted in South Sumatra in January 1997 at a density of 135 palms ha⁻¹.

Data Collection

Fruit bunch yield per palm was recorded for the first six years, from the start of production in June 1999. The cumulative yield of these first six years in tonnes per hectare was used. This was assumed to represent differences between varieties in cumulative yield during the entire economic life of a planting. During the first eight years after field planting, the area of the first fully expanded youngest leaf was annually estimated following the method of Hardon *et al.* (1969). A logistic growth curve was fitted through the mean leaf area values per progeny plot and the asymptotically maximum leaf area (L-max) was estimated as described by Breure and Verdooren (1995).

Estimating the 95% Confidence Interval for the Population Mean of Fresh Fruit Bunch (FFB) Yield

To determine the Upper Confidence Limit (UCL) and Lower Confidence Limit (LCL) of the 95% confidence interval for the population mean FFB in tonnes per hectare of a variety, we calculated the components of variance (VC), *i.e.* a measure of variation, among *pisifera* [(VC (pis.)), among progenies within *pisifera* [VC (prog.)], and between replications [VC (error)]. From these components of variance, the variance of the sample mean [Var (mean)] of the target traits were calculated as:

$$(1/a) \times \text{VC (pis.)} + [1/(a \times b)] \times \text{VC (prog.)} + [1/(a \times b \times r)] \times \text{VC (error)}, \text{ which is}$$

$$\{[1/(a \times b \times r)] \times [b \times r \times \text{VC (pis.)}] + r \times \text{VC (prog.)} + \text{VC (error)}\}$$

where *a* = number *pisifera*, *b* = number of *dura* within *pisifera*, and *r* = number of replications of a set of *a* × *b* progenies.

The LCL and UCL of the population mean FFB in tonnes per hectare with the confidence coefficient of 95% were then calculated as:

$$\begin{aligned} \text{LCL} &= \text{mean} - [\text{SE (mean)} \times 1.96] \\ \text{UCL} &= \text{mean} + [\text{SE (mean)} \times 1.96] \end{aligned}$$

where SE (mean) = the square root of Var (mean) and 1.96 is the right-sided *t* (0.025) point for large samples.

Estimating the 95% Confidence Interval for the Difference of Two Population Means of FFB Yield

The confidence interval of the difference between the population mean of varieties *A* and *B*

was calculated from the standard error (SE) of the difference between the sample means of varieties *A* and *B*, denoted by \bar{y}_A and \bar{y}_B , as:

$$SE(\bar{y}_A - \bar{y}_B) = \text{sqrt}[\text{Var}(\bar{y}_A) + \text{Var}(\bar{y}_B)].$$

where $\text{Var}(\bar{y}_A)$ is Var (sample mean of variety *A*) and $\text{Var}(\bar{y}_B)$ is Var (sample mean of variety *B*).

The LCL and ULC with the confidence coefficient of 95% of the difference between the population means of varieties *A* and *B* were then calculated as:

$$\begin{aligned} \text{LCL} &= (\bar{y}_A - \bar{y}_B) - [\text{SE}(\bar{y}_A - \bar{y}_B) \times 1.96] \\ \text{UCL} &= (\bar{y}_A - \bar{y}_B) + [\text{SE}(\bar{y}_A - \bar{y}_B) \times 1.96] \end{aligned}$$

Power of the Test

When comparing varieties, it is new to include the power of the test, reason why this concept is explained in detail. *Appendix 1* describes, as an example, the procedure of obtaining the minimum number of palms per progeny (sample size) needed to detect a significant superiority for leaf area of progeny *C* against progeny *D*. Testing for significant superiority means first to test the null hypothesis that progenies *C* and *D* have the same leaf area. Then the alternative hypothesis that *C* is superior to *D* with a significant level of 5% is tested. To gain sufficient confidence in the outcome of this test, we need a high probability of at least 80% that the test correctly rejects the null hypothesis and accepts the alternative hypothesis that leaf area progeny *C* is significantly larger than *D*. This acceptance probability is called the power of the test. As shown in *Appendix 1*, power increment can be obtained by increasing the sample size. Furthermore, the chance of rejecting the null-hypothesis can be maximised when the same sample size for the two progenies is used.

Determining the Minimum Number of *Dura* and *Pisifera* Parents in the Sample (*n*) for Detecting Statistical Differences in FFB Yield (t ha⁻¹) between Varieties

We calculated the minimum number of *pisifera* for a given number of *dura* per *pisifera* in the sample, to detect a significant yield difference of $\delta = 2, 3, 4, 5, 6$ and 7 t ha^{-1} between Nigeria *vs.* AVROS and Ekona *vs.* Ghana. The number of *pisifera* involved in the progeny sample was taken from 5 to 11. The calculations were made for two and three replicates. For each *dura* \times *pisifera* progeny, the power of the test (equal or greater than 80%) was calculated. The SAS program PROC POWER was used, with entries of the variance components VC (pis.), VC (prog.) and VC (error) obtained from the database.

The procedure is explained in *Appendix 2* for the interested reader.

Estimating the Optimal Planting Density

The present study assumes that optimal planting density (*D*) is reached at LAI of 5.5. The optimal planting density of the varieties, at average mature leaf size of *L-max*, were then estimated by adopting 40 leaves on the palms, as reported by Gerritsma and Soebagyo (1999):

$$D = (5.5 \times 10\,000) / (\text{L-max} \times 40)$$

RESULTS

Bunch Yield

Inspection of *Table 1* reveals that the variance component of yield among *pisifera* parents [VC (pis)] is clearly highest for Ekona [21.66 *vs.* 0.03 (AVROS), 0.35 (Nigeria) and 5.68 (Ghana)]. This data pattern indicates that there is more variation among *pisifera* in Ekona than in the three other origins. The average cumulative yield (t ha⁻¹) for six years was 101.0 for Ekona, 94.1 for AVROS (lowest yield), 100.9 for Ghana and 103.0 for Nigeria (highest yield) (*Table 1*).

Table 2 presents mean and SE of differences between Nigeria and the three other varieties: AVROS, Ekona and Ghana, respectively. Also, the 95% confidence intervals are given. Note that only the 95% confidence interval for the difference between Nigeria and AVROS does not contain the value zero, which implies that at the 5% level of significance both varieties are unequal with respect to bunch yield (103.4 *vs.* 94.1; *Table 1*).

Effect of the Number of *Dura* Parents Involved in the Progeny Sample on the Length of the 95% Confidence Interval of FFB

Figure 1 shows the half-length [(UCL-LCL)/2] of the 95% confidence interval for differences in mean cumulative yield (t ha⁻¹) between Nigeria and AVROS as a function of number of progenies. The samples involve six *pisifera* with three replicates per progeny. Clearly, the confidence interval sharply decreases with augmenting the number of *dura* from 1 to 5, which results in an increase from 6 (1 \times 6) to 30 (5 \times 6) progenies. When using a still higher number of *dura*, the confidence interval attains a by and large stable level of slight decrement.

Effect of the Number of *Pisifera* Involved in a Fixed Number of Progenies on the Length of the 95% Confidence Interval of FFB

When repeatedly taking a fixed sample of 60 progenies, the half-length (UCL - LCL/2) of the 95% confidence interval decreases sharply with an

TABLE 1. MEAN YIELD VALUES, VARIANCE COMPONENTS AMONG *PISIFERA* [VC (pis.)], AMONG PROGENIES WITHIN *PISIFERA* [VC (prog.)] AND AMONG REPLICATES [VC (error)], ALONG WITH THE STANDARD ERROR (SE) OF THE MEAN AND THE 95% CONFIDENCE INTERVALS OF CUMULATIVE YIELD ($t\ ha^{-1}$) OF FOUR ASD VARIETIES. PLANTING DENSITY ($135\ Palms\ ha^{-1}$), BASED ON $A = 6\ PISIFERA$, $B = 7\ DURA$ PER *PISIFERA* AND $R = 3$ REPLICATES

Variety	Mean	VC (pis.)	VC (prog.)	VC (error)	SE (mean)	95% Confidence intervals		
AVROS	94.1	0.0	3.9	97.4	0.9	92.2	-	95.9
Ekona	101.0	21.7	17.8	128.9	2.2	96.6	-	105.4
Ghana	100.9	5.7	0.0	164.3	1.5	97.9	-	103.8
Nigeria	103.0	0.4	27.8	101.7	1.2	100.6	-	105.5

TABLE 2. MEAN DIFFERENCES IN YIELD ($t\ ha^{-1}$) BETWEEN NIGERIA *vs.* AVROS, EKONA AND GHANA, RESPECTIVELY, ALONG WITH THE STANDARD ERROR (SE) AND 95% CONFIDENCE INTERVALS OF THE DIFFERENCES DERIVED FROM TABLE 1

Pairs of varieties	Mean difference	SE (difference)	95% Confidence intervals		
Nigeria <i>vs.</i> AVROS	9.0	1.5	5.9	-	12.0
Nigeria <i>vs.</i> Ekona	2.0	2.6	-3.0	-	7.0
Nigeria <i>vs.</i> Ghana	2.2	1.9	-1.6	-	6.0

increase of the number of *pisifera* involved from 1 to 4 and then levels off. This trend is clearly visible in Figure 2, which shows the half-length of the 95% confidence interval for differences of mean yield ($t\ ha^{-1}$) between Nigeria and AVROS at varying number of *pisifera* involved in samples of 60 progenies with three replicates. The *pisifera* effect on the width of the confidence interval is of practical importance, since with a fixed number of (60) progenies the same resources are needed for field planting and recording.

Minimum Number of Progenies Required for Detecting a Statistical Difference between Varieties in FFB

Table 3 presents the minimum number of progenies that are needed to obtain a statistical difference in cumulative yield δ from 3 to 7 $t\ ha^{-1}$

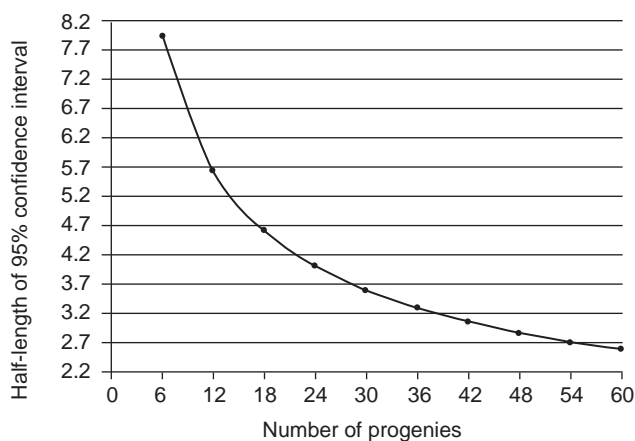


Figure 1. The half-length of the 95% confidence interval for differences of mean cumulative yield ($t\ ha^{-1}$) between Nigeria and AVROS. The samples involve six *pisifera* crossed with one to 10 *dura* resulting in 6 to 60 progenies, tested with three replicates.

between Nigeria and AVROS. The minimum number of progenies consists of either $a = 5, 6, 7$, or 8 *pisifera* crossed with $b = dura$, planted with $r = 2$ or 3 replicates. First of all, we attempt to obtain a difference of $\delta = 4\ t\ ha^{-1}$ with a power of at least 80%. Further, our aim is to use a limit of 60 progenies.

Effect of the Number of *Pisifera*

With about the same power of the test (80%) and by using two replicates, an increase in the number of *pisifera* involved from five to eight can reduce the number of progenies, needed for $\delta = 4\ t\ ha^{-1}$, from 100 to 80.

Effect of the Power of the Test

Let us now assume that we like to test a statistical difference (δ) of 4 $t\ ha^{-1}$ with a power

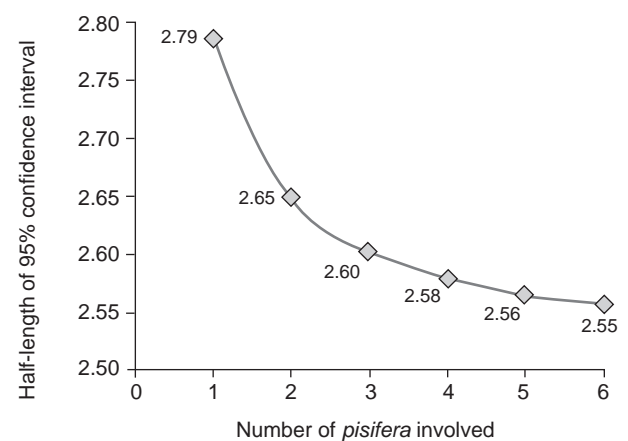


Figure 2. The half-length of the 95% confidence interval for differences of mean yield ($t\ ha^{-1}$) between Nigeria and AVROS as a function of number of *pisifera* (1 to 6) that are involved in samples of 60 progenies, tested with three replicates.

higher than 80%. With, for example, three replicates and involving six *pisifera* in the sample of progenies, an increase in the power of the test from 80.5% to 97.6% would require an increase of progenies from 66 to 114.

Effect of the Number of Replicates

In order to obtain a difference of 4 t ha⁻¹ between Nigeria and AVROS, with a power of the test of 80%, the number of progenies in the sample can be reduced from 100 to 75 by two factors. One is by increasing the number of replicates *r* from 2 to 3. Second is the involvement of five *pisifera* in the sample. To obtain a difference of 5 t ha⁻¹ only 45 progenies are needed. Thus, a difference of 4.5 t ha⁻¹ can be found with 60 progenies [(75 + 45)/2].

Effect of the Variance of the Two Varieties

Table 4 presents the difference in yield between Ekona and Ghana in three replicates. The data pattern reveals that to obtain a yield difference of 7 t ha⁻¹, an increase in the number of *pisifera* involved in the progeny sample from 7 to 9 would result in a reduction of the number of progenies from 112 to a sample of 63.

Note that the statistical difference between Ekona and Ghana (7 t ha⁻¹) clearly exceeds the difference between Nigeria and AVROS (4 t ha⁻¹), while the power of the test (about 80%) as well as the number of progenies (63) are nearly the same.

However, the variance VC (pis.) of Ekona and Ghana (21.7 and 5.7, respectively) are much higher than the values of Nigeria and AVROS (0.0 and 0.4)

TABLE 3. MINIMUM NUMBER OF DURA x PISIFERA PROGENIES, INVOLVING 5, 6, 7 OR 8 PISIFERA REQUIRED TO REACH AN ABSOLUTE YIELD DIFFERENCE BETWEEN NIGERIA AND AVROS (two-sided test at 5% significance level) OF 3 TO 7 t ha⁻¹, ALONG WITH THE STATISTICAL POWER OF THE TEST (%), FROM A PROGENY TEST PLANTED WITH TWO OR THREE REPLICATES

Number of replicates	Statistical power (%)	Yield difference (5 <i>pisifera</i>) ¹					Yield difference (6 <i>pisifera</i>) ¹					Yield difference (7 <i>pisifera</i>) ¹					Yield difference (8 <i>pisifera</i>) ¹				
		3	4	5	6	7	3	4	5	6	7	3	4	5	6	7	3	4	5	6	7
		2	80-89	-	100 ²	60	45	35	162	90	60	42	-	154	84	56	-	-	144	80 ³	56
		-	81.8 ³	80.2	83.4	85.6	80.6	81.4	83.5	84.1	-	81.1	81.1	83.0	-	-	80.3	80.7 ³	84.4	-	-
	90-94	-	135	85	60	45	-	120	78	54	42	210	112	70	49	-	200	112	72	-	-
		-	90.8	91.1	91.9	92.7	-	90.4	91.4	91.6	93.2	90.6	90.3	90.1	90.6	-	90.6	91.4	91.9	-	-
	>95	-	-	105	70	55	-	150	96	66	48	-	140	91	63	49	-	136	88	64	-
		-	-	95.4	95.1	96.3	-	95.2	95.6	95.7	95.7	-	95.1	95.8	95.9	96.8	-	95.4	95.9	96.7	-
3	80-89	135	75 ²	45 ²	35	-	120	66 ²	42	-	-	119	63	-	-	-	112	64	-	-	-
		80.3	81.6 ³	80.1 ³	84.5	-	80.0	80.5 ³	80.9	-	-	82.1	81.0	-	-	-	81.6	83.0	-	-	-
	90-94	-	100	65	45	35	168	90	60	42	-	154	84	56	-	-	152	80	56	-	-
		-	90.3	91.5	91.8	93.3	90.3	90.4	91.9	92.4	-	90.0	90.2	91.7	-	-	90.9	90.1	92.7	-	-
	>95	-	-	80	55	40	-	114 ²	72	48	42	196	105	70	49	-	192	104	64	-	-
		-	-	95.5	95.7	95.8	-	97.6 ³	95.6	95.1	97.7	95.1	95.1	96.1	96.4	-	95.6	95.7	95.4	-	-

Note: ¹Number of *pisifera* involved in the crosses with *dura* palms.

²Number of *dura* x *pisifera* progenies.

³Power of the test (%).

TABLE 4. MINIMUM NUMBER OF DURA x PISIFERA PROGENIES, INVOLVING 7, 8, 9, 10 OR 11 PISIFERA, REQUIRED TO REACH AN ABSOLUTE YIELD DIFFERENCE BETWEEN EKONA AND GHANA (two-sided test at 5% significance level) OF 5 TO 7 t ha⁻¹, ALONG WITH THE STATISTICAL POWER OF THE TESTS (%), FROM A PROGENY TEST PLANTED WITH THREE REPLICATES

Number of replicates	Statistical power (%)	Yield difference (7 <i>pisifera</i>) ¹			Yield difference (8 <i>pisifera</i>) ¹			Yield difference (9 <i>pisifera</i>) ¹			Yield difference (10 <i>pisifera</i>) ¹			Yield difference (10 <i>pisifera</i>) ¹		
		5	6	7	5	6	7	5	6	7	5	6	7	5	6	7
3	80-89	-	-	112 ²	-	-	72	-	144	63 ²	-	100	-	-	99	-
		-	-	80.1 ³	-	-	81.2	-	80.4	83.4 ³	-	80.9	-	-	83.9	-
	90-94	-	-	-	-	-	-	-	-	135	-	-	90	-	242	77
		-	-	-	-	-	-	-	-	90.3	-	-	90.2	-	90.3	90.9
	>95	-	-	-	-	-	-	-	-	-	-	-	240	-	-	143
		-	-	-	-	-	-	-	-	-	-	-	95.1	-	-	95.1

Note: ¹Number of *pisifera* involved in the crosses with *dura* palms.

²Number of *dura* x *pisifera* progenies.

³Power of the test (%).

(Table 1). It should be noted that, considering the negligible difference in yield between Ekona and Ghana (101.0 vs. 100.9, Table 1), only an unrealistically high number of progenies would detect a significant difference between these two varieties.

We conclude that, considering the practical number of progenies that fit in a plantation unit, it seems that a sample of 60 progenies is the most acceptable option for the test. Depending on the variance component of the *pisifera* parent, a difference of 4.5 to 7 t ha⁻¹ can be obtained, which is equivalent to a difference of about 4.5% to 7% between varieties of our tests.

Maximum Leaf Area and Optimal Planting Density

Table 5 presents the means of the maximum leaf area (L-max) of the four varieties under study. For each variety also the 95% confidence interval is given.

Clearly, the L-max was highest for Ekona (10.14 m²) and lowest for Ghana (7.83 m²). Moreover, the 95% confidence interval for Ghana (7.30 – 8.37) did not show any overlap with those of the other three varieties.

Table 6 presents, in paired comparisons of the four varieties, the difference in L-max as well as the SE and the 95% confidence intervals of the differences.

When looking at the difference between the paired varieties (Table 6), the confidence intervals for Ghana vs. AVROS, Ekona and Nigeria, respectively, do not contain the value zero. In other words, Ghana vs. AVROS, Ekona and Nigeria are dissimilar in maximum leaf area at a level of 5% significance.

On the other hand, AVROS vs. Ekona and Nigeria, respectively, and Ekona vs. Nigeria do contain the value zero, which implies that the hypothesis of equality of maximum leaf area among these pairs of varieties is not rejected.

By applying the mean L-max values of the four varieties (Table 4), the optimal planting densities (palms ha⁻¹), as estimated from the Equation $D = (5.5 \times 10\,000) / (L\text{-max} \times 40)$, become 144, 136, 143, and 176 for AVROS, Ekona, Nigeria and Ghana, respectively. If we only consider the practical range of planting densities and take into account the range of L-max values of the 95% confidence intervals, we conclude that Ghana can be planted at 160 palms ha⁻¹ (corresponding to L-max of 8.59 m²), vs. the three other varieties at 143 palms ha⁻¹ (corresponding to L-max of 9.62 m²).

DISCUSSION

Number of Progeny Samples

The present study shows that records of at least 60 individual progenies, preferably involving at least six *pisifera*, are necessary when reliably estimating the mean of all *dura* × *pisifera* progenies of a variety. A sufficient number of *pisifera* is feasible when all selected *pisifera*, as occurring for example in *tenera* selfings of La Me or Yangambi varieties in the CIRAD system of breeding (Baudouin, 1997), are used as male parents for seed production. However, a connected crossing design that allows estimating the breeding value of individual *pisifera*, as reported by Breure and Verdooren (1995), permits selecting

TABLE 5. MEANS OF THE MAXIMUM LEAF AREA (L-max), TOGETHER WITH THE VARIANCE COMPONENTS AMONG *PISIFERA* [VC (pis.)], AMONG PROGENIES WITHIN *PISIFERA* [VC (prog.)] AND AMONG REPLICATES VC (error). THE STANDARD ERROR (SE) AND THE 95% CONFIDENCE INTERVALS ARE GIVEN FOR EACH VARIETY. THE DATA PATTERNS ARE ASSEMBLED FOR A = 6 *PISIFERA*, B = 7 *DURA* PER *PISIFERA* AND R = 3 REPLICATES

Variety	Mean	VC (pis.)	VC (prog.)	VC (error)	SE (mean)	95% Confidence intervals		
AVROS	9.57	0.088	0.144	0.510	0.149	9.28	-	9.86
Ekona	10.14	0.665	0.631	0.804	0.364	9.43	-	10.86
Ghana	7.83	0.376	0.437	0.255	0.274	7.30	-	8.37
Nigeria	9.62	0.271	0.855	0.568	0.265	9.10	-	10.14

TABLE 6. DIFFERENCES BETWEEN THE MEANS OF THE MAXIMUM LEAF AREA (L-max) FOR PAIRS OF VARIETIES, ALONG WITH THE STANDARD ERROR (SE) AND 95% CONFIDENCE INTERVALS OF THE DIFFERENCES IN MEANS DERIVED FROM TABLE 5

Pairs of varieties	Mean (difference)	SE (difference)	95% Confidence intervals		
Ekona vs. AVROS	0.57	0.39	-0.20	-	1.34
Nigeria vs. AVROS	0.05	0.30	-0.55	-	0.64
Ekona vs. Nigeria	0.52	0.45	-0.36	-	1.405
AVROS vs. Ghana	1.74	0.31	1.13	-	2.35
Ekona vs. Ghana	2.31	0.46	1.42	-	3.20
Nigeria vs. Ghana	1.79	0.38	1.04	-	2.53

a smaller number of *pisifera* per variety. In that case an excessive number of progenies are needed to obtain reliable statistical differences between varieties, which renders the tests too laborious. A way out is to involve the entire group of selected *pisifera* parents of all varieties provided by a seed agency. A final choice of individual varieties could then be based on the breeding values of yield as well as other traits of the *pisifera* parent(s) involved. One may argue that plantation companies only want to compare different seed agencies anyway. Sometimes, however, as is the case in Indonesia, licenses for seed sales are granted for specific varieties. Or it happens that typical characteristics of the *pisifera* (involved) are used for sales promotion.

Asking agencies to prepare a sample for a comparative trial well in advance of the start of testing is risky, since some agencies may select the best crosses. An option could be to draw a random sample of all crosses that are stored (available) by applying the Excel function *Random Permutation*.

Testing Trials

The need to test oil palm varieties may be questioned. The reason is that by the time results of comparing planting material from the seed agencies are known, the outcomes do not apply anymore to a later generation. However, one may assume that progenies still came from the same distinct origins, and also that parent selection of seed agencies had advanced at about the same extent. Note that, in practice, plantation companies do compare material on the basis of specific seed agencies.

Plantation companies may lack the experience or facilities to record data from individual palms of 60 progenies per variety. A convenient way to facilitate data assembling is to plant the progenies in plots of half-rows, which allows weighing the entire number of bunches of one plot on the road side by using a simple tripod, as described by Breure and Verdooren (1995). To avoid the inclusion of bunches from replacement or abnormal palms, these bunches may be frequently marked while still on the palms and separated from the bulk of bunches per plot before weighing takes place.

Selection Traits

It could be argued that when palms are planted in half-rows the variation in competition for light between neighbouring plots would bias the yield. However, as Nazeeb *et al.* (2007) showed, the effect of planting density due to light competition commences not until palms are about five years in production. By that time usually all records of the trials will be obtained.

In order to estimate optimal planting density, it proves useful to determine leaf area at maximum

crown size. The varieties are then estimated for the expected yield per hectare, instead of the yield outcome from the actual planting density of the experiment. Leaf area measurements also serves as a guide for spacing the palms of selected varieties. Measuring the youngest leaf at annual intervals to extrapolate the asymptotically maximum leaf area (L-max), as Breure applied in his 2010 article, may not be feasible for estate recording, though. A reliable alternative forms the measurements of a few young leaves at the end of the recording period. At that time, which is about seven years after planting, the crown size approaches its maximum level (Breure, 2010). As estates are already familiar with identifying frond 17 for leaf analysis, which is the third leaf in the spiral of the youngest frond, the measurement of frond 9 (above frond 17 in the same spiral), along with the youngest frond in other two specified spirals would be a convenient option in order to estimate L-max.

In the present study, the yield of Nigeria and Ghana varieties (103.0 *vs.* 100.9 t ha⁻¹) was by and large similar (Table 1), but the estimates of optimal practical densities were clearly unequal (143 *vs.* 160 palms ha⁻¹). For a final selection, when adjusted to the yield per hectare at optimal planting densities for both varieties, it is necessary to check whether the test of equality (H_0), with a power of 80%, is rejected or not. Such an extrapolation of yield per palm obtained at a standard to a (slightly) discrepant planting density is justified. The reason is that the yield per palm during the first five years is virtually still independent of planting density, as pointed out above.

Comparing Oil Palm Varieties

Corley and Tinker (2016) reviewed trials that compare origins or varieties [for example, Asmono *et al.* (1999); Sterling and Alvarado (1995); Lee (1999)]. None of these trials, however, include the SE of means, while usually the number of progenies under test is too small for an adequate statistical backing.

Moreover, the conventional interpretation of the results does not take into consideration the power of the test. Note that in a paired comparison of two varieties the 95% confidence interval of the difference in mean is equivalent to a test of zero difference at a significance level of 5% with a power of only 50%. In other words, 50 out of 100 cases the test rejects a false null hypothesis. Power increment is the statistical procedure to minimise the probability of rejecting a false null hypothesis, as applied in the presentation of outcomes of the test in Tables 3 and 4.

These tables present the statistical comparisons of progenies for two sets of parents (Nigeria *vs.*

AVROS and Ghana *vs.* Ekona). Further, a certain number of replicates are taken in the analysis, which, along with increasing the size of the sample of the test, results in enlarging the power of the test to 80% and even more. Furthermore, in contrast to confidence intervals with a fixed power of 50%, for each *dura* × *pisifera* cross of the test, the power of the difference in yield between the varieties can be evaluated. Table 3 shows, as an example, a test with a sample of 60 progenies involving five *pisifera* with two replicates. A difference between Nigeria and AVROS of 5 t ha⁻¹ is statistically significant with a power of 80.2%, while differences of 6 and 7 t ha⁻¹ are significant with a power of 91.9% and more than 95%, respectively. Including the power of the test when comparing varieties, with a certain sample size and number of replicates, therefore, allows a more reliable choice of varieties than with conventional variety tests.

The higher VC (pis.) of Ekona implies that there is more scope for selecting individual Ekona *pisifera* parents than for *pisifera* of the other varieties. A caveat of our study is the specific group of varieties under test. Obviously, data patterns found elsewhere, for example, in another set of varieties from other seed agencies, may generate results that considerably deviate from our comparisons. Still the outcomes, as presented in Tables 3 and 4, may serve as a guide for setting up new test trials. In particular, the number of progenies and *pisifera* involved in the samples are of relevance. As a rule of thumb, we recommend to draw samples of 60 progenies, originated from, for example, six *pisifera* and 10 *dura* per *pisifera*.

Another caveat is that this study is based on results of only one location, ignoring a possible variety × site interaction. For a final comparison, identical trials at distinct environments are needed and its implication on power analysis, if any, determined.

CONCLUSION

Plantation companies and government organisations tend to compare oil palm varieties based on yield per hectare obtained at one, fixed, density. The potential yield of a variety can be more reliably obtained from the yield per palm, along with the optimal planting density, the latter estimated from the leaf area at mature crown size.

The present study defines the 95% confidence intervals of paired comparisons of four ASD varieties (Nigeria, AVROS, Ghana and Ekona). By enlarging the size of the samples till they contain at least 30 *dura* × *pisifera* progenies and involving at least four *pisifera*, the 95% confidence intervals of yield strongly decreases, implying that the comparison test proportionally gains in accuracy. The width of the confidence interval levels off in case of a higher number of progenies as well as *pisifera* involved.

The 95% confidence interval is equivalent to a test of differences in yield with a power of only 50%, which is the probability that the test correctly rejects a false null-hypothesis of zero difference between two varieties. The present study therefore estimates the significant yield difference that can be obtained between (re-sampled) progeny sets of paired varieties at a power equal or higher than 80%. With a power of the test of about 80% and the use of three replicates, depending on the variance component of the *pisifera* parents, a sample of 60 progenies is sufficient in order to find a significant (at 5% level) difference in yield of 4.5% between Nigeria and AVROS and 7% between Ghana and Ekona, when six and nine *pisifera*, respectively, are involved in the sample.

To allow bulk recording of total yield per progeny plot, unbiased yield records can be obtained when planting the samples in plots of half-rows, as described by Breure and Verdooren (1995). To ensure a sufficient number of crosses that reliably represent the entire population of progenies within a variety, it is desirable to take random progeny samples of at least 12 *dura* per *pisifera* and involving five *pisifera*.

ACKNOWLEDGEMENT

The permission of PT Selapan Jaya to publish the results and the assistance of the staff of Bina Sawit Makmur in the collection of the data are gratefully acknowledged. Special thanks go to L R Verdooren whose suggestions and statistical analysis made a vital contribution; the article owes much to his help. We wish to thank M Mustiqa Siregar for his assistance in the statistical analysis, R H V Corley for his comments and P J G Keuss for editorial assistance.

REFERENCES

- ASMONO, D; IKWAN, M and PURBA, A R (1999). Genetic improvement of IOPRI oil palm using *pisifera* derived from Congo germplasm. *Proc. of the 1999 PORIM International Palm Oil Conference*. PORIM, Bangi. p. 351-360.
- BAUDOQUIN, L; BARIL, C; CLEMENTE-DEMANGE A; LEROY, T and PAULIN, D (1997). Recurrent selection of tropical tree crops. *Euphytica*, 96: 101-114.
- BREURE, C J (1985). Relevant factors associated with crown expansion in oil palm (*Elaeis guineensis* Jacq.). *Euphytica*, 34: 161-175.
- BREURE, C J (2003). The search for yield in oil palm: basic principles. *The Oil Palm, Management for Large*

and Sustainable Yields (Fairhurst, T H and Härdter, R eds.). Potash & Phosphate Institute/Potash Institute of Canada and International Potash Institute. p. 58-99.

BREURE, C J (2006). Performance of ASD's oil palm parent material in South Sumatra: the search for elite planting material for Indonesia. *ASD Oil Palm Papers*, 29: 19-30.

BREURE, C J (2010). Rate of leaf expansion: a criterion for identifying oil palm (*Elaeis guineensis* Jacq.) types suitable for planting at higher densities. *NJAS – Wageningen J. Life Sciences*, 57: 141-147.

BREURE, C J and VERDOOREN, L R (1995). Guidelines for testing and selecting parent palms in oil palm. Practical aspects and statistical methods. *ASD Oil Palm Papers*, 9: 1-68.

CORLEY, R H V and TINKER, P B (2016). *The Oil Palm*. 5th ed. Blackwell Science, Oxford, UK.

GERRITSMA, W and SOEBAGYO, F X (1999). Analysis of the growth of leaf area of oil palm in Indonesia. *Experimental Agriculture*, 35: 293-308.

HARDON, J J; WILLIAMS, C N and WATSON, I (1969). Leaf area and yield in the oil palm in Malaysia. *Experimental Agriculture*, 5: 25-32.

ISA, Z A; KUSHAIRI RAFIL, A M Y; SALEH, G and RAJANAIDU, N (2009). Variation in FFB and yield components in Malaysian oil palm (*Elaeis guineensis* Jacq.) DxP planting materials under various

planting densities and their correlations with frond production, rachis length and height. *Proc. of the International Palm Oil Congress - Agriculture, Biotechnology and Sustainability Conference*. Vol. II. MPOB, Bangi. p. 700-736.

LEE, C H (1999). Yield potential of Golden Hope DxP oil palm planting materials. *Proc. of the Seminar 1996: Sourcing of Oil Palm Planting Materials for Local and Overseas's Joint Ventures* (N Rajanaidu and Jalani, B S eds.). PORIM, Bangi. p. 106-117.

NAZEEB, M; TANG, M K; LOONG, S G and SYED, S B (2007). Variable density plantings for oil palms (*Elaeis guineensis* Jacq.). *Proc. of the PIPOC 2007 International Palm Oil Congress – Agriculture, Biotechnology and Sustainability Conference*. p. 127-167.

STERLING, F and ALVARADO, A (1995). Ekona and Calibar alternative sources of male parents in the commercial production of the oil palm seed. *ASD Oil Palm Papers*, 11: 23-32.

TURNER, P B and GILLBANKS, R A (2003). *Oil Palm Cultivation and Management*. Incorporated Society of Planters, Kuala Lumpur.

VON UEXKULL, H; HENSON, I E and FAIRHURST, T (2003). Canopy management to optimize yield. *The Oil Palm, Management for Large and Sustainable Yields* (Fairhurst, T H and Härdter, R eds.). Potash & Phosphate Institute/Potash Institute of Canada and International Potash Institute. p. 163-180.

BASIC CONCEPT FOR ESTIMATING STATISTICAL SUPERIORITY BETWEEN TWO PROGENIES

To illustrate the idea of the power of a test, we first consider determining the minimum number of palms to detect a certain statistical difference between two progenies, for example, whether the mean leaf area of progeny (*dura* x *pisifera* cross) C is δ m² higher than the leaf area of progeny D. To test whether the leaf area of progeny C is significantly larger than progeny D, we assume that the leaf area is normally distributed with unknown population means μ_C and μ_D and known variances of leaf area per palm of σ_C^2 and σ_D^2 (in practice, the variances must of course be first estimated and then we need to use the power of the t-test). For progenies of oil palm, it is reasonable to assume that $\sigma_C^2 = \sigma_D^2 = \sigma^2$. Testing for significant superiority means statistically testing first the null hypothesis (H_0) that C and D have the same leaf area: ($\mu_C - \mu_D = 0$). Then, the alternative hypothesis (H_a) that C is superior to D: ($\mu_C - \mu_D > 0$) is tested. In this section, we take a significance level of $\alpha = 0.05$.

Let us assume that there are n_C palms available for progeny C and n_D palms for progeny D, where $n_C + n_D = N$ and, hence, $n_D = N - n_C$ and the sample means of leaf area are denoted as \bar{y}_C and \bar{y}_D . The test statistic that we use for this test is

$$T = [(\bar{y}_C - \bar{y}_D) - (\mu_C - \mu_D)_{H_0}] / \sqrt{\text{var}(\bar{y}_C - \bar{y}_D)}, \text{ which is}$$

$$T = [(\bar{y}_C - \bar{y}_D) - 0] / \sqrt{\sigma^2 [1/n_C + 1/(N - n_C)]}.$$

T is under H_0 normally distributed with mean 0 and variance 1, which is called the standard normal distribution Z . The table of Z distribution shows that at a significance level of $\alpha = 0.05$, the critical value is 1.645 or the probability (Pr) that $Z > 1.645 = 0.05$. If the outcome of T falls in the critical region ($T > 1.645$), we reject the null-hypothesis H_0 and accept the alternative hypothesis H_a and, hence, conclude that progeny C is superior to progeny D.

Power of the Test

To be quite sure to reject the null hypothesis H_0 that C and D have the same leaf area: ($\mu_C - \mu_D = 0$), and accept the alternative hypothesis H_a that C differs δ m² from D: ($\mu_C - \mu_D = \delta > 0$) we use a probability of at least 80%. In other words, we use a probability of at least 80% that the outcome of T under the alternative distribution H_a falls in the rejection region of H_0 . This acceptance probability

is called the power of the test. We can increase the power of the test by increasing the total sample size $N = 2n$, where each progeny has n palms.

Minimum Sample of Palms per Progeny for Detecting Significant Differences.

With the same sample means ($\bar{y}_C - \bar{y}_D$), the test-statistic T is largest if the denominator is minimal which is the case when $n_C = n_D = n = N/2$. So the test has the highest chance to reject the null-hypothesis if we take a sample with the same number of palms per progeny: $n_C = n_D = n = N/2$.

The question is now that we like to determine a minimal progeny sample n that is needed to detect a statistical superiority between population means ($\mu_C - \mu_D$) of progeny C and D of δ m² with a power of at least 80%.

Under the alternative hypothesis (H_a): $\mu_C - \mu_D = \delta > 0$ the test-statistic T is now:

$$[(\bar{y}_C - \bar{y}_D) - \delta] / \sqrt{\sigma^2 [1/n + 1/n]}, \text{ which is}$$

$$T = Z - \delta / \sqrt{\sigma^2 (2/n)}$$

For a power of the test of at least 0.80, we have under H_a that the probability: $\text{Pr}(T \geq 1.645) = 0.80$ or

$$\text{Pr}(Z - \delta / \sqrt{\sigma^2 (2/n)} \geq 1.645) = \text{Pr}[Z \geq 1.645 + \delta / \sqrt{\sigma^2 (2/n)}] = 0.80.$$

The table of Z shows that for a probability of at least 0.80 the Z score must be:

$$\geq -0.842 \text{ or } 1.645 + \delta / \sqrt{\sigma^2 (2/n)} \geq -0.842.$$

$$\text{Hence, } 1.645 + 0.842 \geq -\delta / \sqrt{\sigma^2 (2/n)} \text{ or } (1.645 + 0.842)^2 / [(-\delta)^2 / \sigma^2] \leq n/2.$$

The minimum number of palms of each progeny (n) for a test with a power of 80% to detect a superiority of C against D of δ m² with a significance level of $\alpha = 0.05$ is then:

$$n \geq 2 (1.645 + 0.842)^2 \times (\sigma^2 / \delta^2).$$

The Z scores of 90% and 95% are 1.28 and 1.65, respectively. Note that for a power of 50% the Z score becomes zero. So the number of palms (sample size) for obtaining the significant superiority of progeny C is directly related to the power of the test.

DETERMINING THE MINIMUM NUMBER OF *DURA* AND *PISIFERA* PARENTS INVOLVED IN THE PROGENY SAMPLE (n) FOR DETECTING STATISTICAL DIFFERENCES IN FFB YIELD ($t\ ha^{-1}$) BETWEEN VARIETIES

The objective of the present article is to determine the difference in FFB ($t\ ha^{-1}$) of variety *A* against *B* and, hence, we need to determine the minimum number of progenies for detecting a significant difference δ between variety *A* and *B*. For this calculation, we need to know the variance components VC (pis.), VC (prog.) and VC (error), which were determined from the database of this article.

Let us assume that a plantation company likes to find a difference between varieties offered by various seed agencies and aims at a difference of at least $\delta\ t\ ha^{-1}$. To determine the progeny sample (n) needed, we must calculate the minimum number of progeny plots of the test ($n = a \times b \times r$) to detect a significant difference δ between *A* and *B*, where a = number of *pisifera* for a given number b = number of *dura* per *pisifera* and r = number of replicates. In other words, we want to determine the minimum number of *total* progeny plots of the test that are needed, to reject the null hypothesis $H_0: (\mu_A - \mu_B) = 0$ and accept the alternative hypothesis $H_a: |\mu_A - \mu_B| > \delta\ t\ ha^{-1}$, with $\delta > 0$, at a significance level $\alpha = 0.05$ and a power of at least 80%. Because we consider either $\mu_A - \mu_B = \delta$ or $\mu_B - \mu_A = \delta$, we need a *two-sided critical region* at a significance level $\alpha = 0.05$. The H_0 is then rejected if the test-statistic T falls in the region $T \leq -1.960$ or $T \geq +1.960$, because the probability that $Z \leq -1.960$ is 0.025 and the probability that $Z \geq +1.960$ is 0.025.

The minimum number of total progeny plots ($n = a \times b \times r$) for each variety of the test was calculated by using the SAS program PROC POWER as follows.

We insert in the formula of Var (mean) a given number r of replicates and b of *dura* per *pisifera* for each variety *A* and *B* and insert $SE(\bar{y}_A) = \text{sqrt}[\text{Var}(\bar{y}_A)]$ with the variance component of variety *A* and of $a = 1$:

$$\text{Var (mean)} = [1/(a \times b \times r)] \times [b \times r \times \text{VC (pis.)} + r \times \text{VC (prog.)} + \text{VC (error)}]$$

Analogous we find $SE(\bar{y}_B)$ with the formula.

The SAS program PROC POWER gives then the minimum number of *a pisifera* for each variety *A* and *B* for a given number b of *dura* per *pisifera* with a given number r of replicates plots per progeny. Doing this for another number b of *dura* per *pisifera* with a given number r of replicates plots per progeny, we again determine the minimum number of *a pisifera*.

From the minimum number of progeny plots ($n = a \times b \times r$), determined with PROC POWER for a power equal or greater than 80%, we then derive the minimum number b of *dura* crossed with a number of *a pisifera* ranging from 5 to 11 to detect a significant difference δ between *A* and *B* of 2, 3, 4, 5, 6 and 7 $t\ ha^{-1}$ between varieties *A* and *B* for r is two or three replicates.