SOME CONSIDERATIONS IN ANALYSING THE YIELD DATA OF COCONUT*

JACOB MATHEW AND K. VIJAYA KUMAR
Central Plantation Crops Research Institute,
Kasaragod-670 124, Kerala

ABSTRACT

Examination of yield data of coconut for individual years revealed that the distribution is always positively skewed and leptokurtic. Pooling the data for consecutive years as well as consideration of palms in groups did not improve the distribution. Among the different transformations tried, $\sqrt{x \times 10}$ transformation was found to bring the data to a near normal distribution in most cases. Another difficulty experienced while analysing the yield data of coconut experiments is the large error variance. With pooling of data for consecutive years. variance was found to get reduced. Covariance analysis was found to result in increased efficiency.

INTRODUCTION

Though the statistical techniques followed in experimentation with annual crops and perennial crops are essentially the same, one specific difficulty experienced while analysing the yield data of coconut experiments is the large error variance. This is mainly due to the heterogeneous nature of the experimental material, differential response of individual plants to the varying weather conditions over the years, and soil fertility differences arising out of the vast area required for laying out any meaningful experiments. Because of the large error variance, real treatment differences are often masked and go undetected. Reduction in residual variation is generally achieved through methods of calibration and covariance analysis. Shrikhande (1957) and Abeywardena (1970) have illustrated the use of pre-experimental yield data while designing new experiments with coconuts. Govinda Iyer (1957) has observed that the pre-experimental data for three years as the concommittant variate will help in detecting treatment differences most efficiently. According

*CPCRI Contribution No. 257.

ing to Abraham and Kulkarni (1963), data collected about two years immediately prior to the experimental period was sufficient for covariance analysis. In the light of the above, the variability in the yield of coconuts in the different years, its distribution, relationship between yields obtained in different years, and ways to reduce the experimental error, have been studied and results are presented.

Materials and Methods

Individual tree yield data available for the period 1961–80 from two plots at Central Plantation Crops Research Institute, Kasaragod, was made use of for this study. The palms in Plot 1 (RS 28 Central Block, n = 155 palms) were 35-40 years old at the beginning of this period, and those in Plot-2 (RS 29 North Block, n = 354 palms) were aged 45-50 years. The palms were growing under rainfed conditions and receiving the recommended dose of fertilizers. Normality of distribution of yield data was tested by working out the coefficients for skewness (b_1) and kurtosis (b_2). The variances for individual years and for cumulative yields for five years, and the reduction in variance due to covariance analysis using one to five years' pre-experimental yield data, were also worked out.

Results and Discussion

The valid application of tests of significance in analysis of variance requires that experimental errors should be independently and normally distributed with a common variance (Eisenhart, 1947). This, in other words means that in an experiment-free material, the distribution of the original data should be normal. Tests for normality revealed that in general, the distribution of annual yield of coconut is always positively skewed and leptokurtic, generally following a Pearson Type IV distribution (Table 1). This is due to the presence of proportionally higher percentage of palms giving below average yield. In the years of low production skewness was found to be more. Among the two plots considered, coefficient of skewness was comparatively less in Plot 2, where the yield was higher. Similarly, distribution was found to be leptokurtic in the years where yield was generally lower. Studies conducted in other perennial crops like cocoa, coffee and oil palm (Butters, 1964; Chapas, 1961; Cunningham and Burridge, 1957; Lotode and Muller,

1974) have also shown that the distribution of their yield data is far from normal.

Table 1. Measures of central tendency (mean), skewness (b_1) and kurtosis (b_2,) for the yield data of coconut at Kasaragod

Plot 1 RS 28 Central Block (n=155)				Plot 2 RS 29 North Block (n=354)			
Year	Mean	b_1	b_2	Year	Mean	b_1	b_2
1967	48.8	0.98**	4.41*	1961	56.2	0.19	2.93
1968	44.2	0.83**	3.61	1962	54.5	1.51**	9.92**
1969	40.2	0.88**	3.92*	1963	47.2	0.63**	4.19**
1970	53.6	0.78**	3.26	1964	66.6	1.00**	5.85**
1971	47.6	1.41**	7.43**	1965	44.5	0.56**	3.82**
1972	58.2	0.73**	3.64*	1966	68.6	0.96**	5.00**
1973	50.1	0.95**	4.35**	1967	58.3	0.40**	3.99**
1974	41.4	0.97**	4.09*	1968	53.8	0.74**	4.07**
1975	55.0	0.52**	3.26	1969	59.1	1.99**	17.68**
1976	43.4	0.81**	4.05*	1970	69.0	0.53**	4.33**

*Significant at 5% level, **Significant at 1% level

Among the different transformations tried, square root transformation of the form $\sqrt{x+10}$ was found to be the best. This was effective in reducing the magnitude of skewness and kurtosis in

Table 2. Effect of square-root transformations on the normality of distribution of yield data of coconut

Plot 1:RS 28 Central Block (n=155)					Plot 2:RS 29 North Block (n=354)				
	$\sqrt{x+3/8}$ transformation		$\sqrt{x+10}$ transformation			$\sqrt{x+3/8}$ transformation		$\sqrt{x+10}$ transformation	
Year	b_1	b_2	b_1	b_2	Year	b_1	b_2	b_1	b_2
1967	–0.03	2.93	0.22	2.88	1961	–0.67**	3.58*	–0.44**	3.09
1968	–0.16	2.61	0.05	2.54	1962	–0.02	4.62**	–0.29*	4.67**
1969	–0.18	2.98	0.16	2.79	1963	–0.31**	3.30	–0.08	3.09
1970	–0.06	2.78	0.16	2.55	1964	–0.13	4.23**	–0.09	3.96**
1971	–0.09	4.29**	0.33*	4.07*	1965	–0.43**	3.15*	–0.14	2.85
1972	–0.47**	3.73*	–0.06	3.06	1966	–0.06	3.62*	–0.18	3.59
1973	–0.12	3.21	0.18	3.01	1967	–0.41**	3.33	–0.26*	3.21
1974	–0.17	2.97	0.20	2.84	1968	–0.15	3.14	–0.05	2.96
1975	–0.09	2.70	–0.01	2.51	1969	–0.13	5.47**	–0.20	5.61**
1976	–0.12	3.04	0.13	2.88	1970	–0.48**	4.11**	–0.28*	3.75**

*Significant at 5% level **Significant at 1% level.

all the years. In some cases, a shift to negative skewness was also noticed. The conventional types of square root transformations like $\sqrt{x}$, $\sqrt{x+3/8}$, and $\sqrt{x}+\sqrt{x+1}$ were found to be of limited use only, in the case of coconut (Table 2).

The effect of pooling data for consecutive years on the distribution was further examined. Since bienniality has been shown to be a significant feature in coconut also (Abeywardena, 1962, and Jacob Mathew, unpublished), the effect of extreme values obtained in alternate years gets evened out when the data for adjoining years are pooled. The coefficients of skewness and kurtosis were found to get reduced when data for two to five years were pooled. With transformation, skewness was found to decrease, accompanied by a shift from positive to negative skewness (Table 3).

Table 3. Effect of pooling data for consecutive years on the normality of distribution of yield of coconut (Plot 2:RS 29, n=252)

Period	Original data		$\sqrt{x+3/8}$ transformation		$\sqrt{x+10}$ transformation	
	b_1	b_2	b_1	b_2	b_1	b_2
1961–62	0.39**	3.65*	−0.34*	3.68*	−0.21	3.53*
1962–63	0.57**	4.17**	−0.25	3.89**	−0.09	3.73*
1961–63	0.34*	3.58*	−0.37**	3.63*	−0.24	3.47
1962–64	0.59**	4.03**	−0.15	3.74*	−0.02	3.61*
1961–64	0.41**	3.62*	−0.27*	3.60*	−0.15	3.46
1962–65	0.39**	3.58*	−0.31*	3.70*	−0.17	3.48
1961–65	0.30*	3.47	−0.38**	3.64*	−0.25	3.45
1962–66	0.52**	3.76*	−0.16	3.71*	−0.04	3.54*

Cunningham and Burridge (1959) in cocoa, and Chapas (1961) in oil palm could get normal distribution of yield data by grouping of trees. In coconut also, grouping of data for trees was found to reduce the skewness (Table 4). In the different years for which this was attempted, a gradual reduction in skewness was noticed with increase in plot size, and with square root transformation of the form $\sqrt{x+10}$, the distribution of yield was found to become normal. However, in years where the distribution of original data itself was highly skewed, not much improvement could be noticed with grouping of trees, even with square-root transformation (Table 4).

Table 4. Effect of multi-tree plots on the normality of distribution of yield of coconut (Plot 2:RS 29, n=354)

Year	Plot size (No. of trees)	Original data		$\sqrt{x+10}$ transformation	
		b_1	b_2	b_1	b_2
1967	1	0.41**	4.04**	-0.26*	3.21
	2	0.40*	3.97*	-0.08	3.60
	3	0.45*	3.03	0.16	2.79
	4	0.56*	3.12	0.34	2.82
	5	0.64*	3.07	0.46	2.83
	6	0.45	2.85	0.27	2.69
1968	1	0.74**	4.07**	0.05	2.97
	2	0.63**	2.54	0.21	2.95
	3	0.30	2.67	0.01	2.43
	4	0.43	2.70	0.19	2.65
	5	0.32	2.62	0.13	2.52
	6	0.70*	3.76	0.41	3.24
1969	1	1.99**	17.68**	0.20	5.61**
	2	1.60**	9.81**	0.73**	5.66**
	3	1.46**	7.24**	0.91**	4.95**
	4	1.19**	5.91**	0.70**	4.71**
	5	1.26**	5.28**	0.93**	4.34**
	6	1.17**	4.84	0.87**	4.25*
1970	1	0.53**	4.33**	-0.28*	3.75**
	2	0.73**	4.42	0.20	4.01*
	3	0.42	2.73	0.21	2.56
	4	0.50*	3.29	0.24	3.01
	5	0.40	3.54	0.14	3.33
	6	0.27	2.39	0.13	2.36

*Significant at 5% level **Significant at 1% level.

It follows from the above analysis that square-root transformation of the form $\sqrt{x+10}$ is necessary while analysing the yield data of coconut, for normalising the distribution. Even when multi-tree plots are used, the above transformation is suggested before the yield data is subjected to tests of significance, using analysis of variance technique.

The inter-relationship between the yields obtained in the different years was also examined. Coefficients of correlation, r_{piq}, were calculated for the mean yields of si gle trees between an earlier period of p (=0, 1, 2, 3, 4 and 5) years and a later period of q (=1, 2, 3, 4, and 5) years separated by an interval of i (=0, 1, 2, 3, 4 and 5) years. The yield data for 1961–80, for a group of 252 palms (RS-29, North Block), was used for this study. It was seen that, though the yields obtained in the different years are highly correlated, the relationship is comparatively weak when the annual data for immediately preceding and succeeding years are considered. This is due to the alternate bearing tendency shown by some of the palms. Compared to this, the correlation was much higher when there was a gap of one year, between the two years under consideration. When the data for groups of years were considered, the coefficients of correlation were found to go upto 0.9. However, marginal decreases in values were noticed when the gap separating the earlier and later periods was increasing.

The use of ancillary variables in the analysis of covariance for reduction of expe imental errors is a widely accepted practice. The variation in perennial plants is mainly due to environmental variation and also inherent genetic variability. Covariance analysis is intended to even out the ge etic variation, which is generally of greater magnitude. Shrikhande (*loc. cit*). has found that genetic and environmental components are in the ratio of 2:1 or 3:2, though Pankajakshan (1960) later showed that environmental component is more important when data are considered in blocks of four years or more. In view of the close relationship observed betw en the yields obtained in the different years, covariance analysis was attempted using different combinations of earlier and later periods. Relative efficiency of this type of analysis, as measured by the ratio of inverses of error variances, was calculated in all these cases (Table 5). Pooling the data for consecutive years was found to increase the efficiency by about 50 to 60 per cent. However, when pre-experimental data were used in covariance analysis, substantial reduction in variance and consequent improvement in efficiency was noticed. Use of two years pre-treatment data was found to almost double the efficiency, compared to what was obtained with single year's data. Relative efficiency was seven times higher with the use of two years pre-treatment data and nine

Table 5. Efficiency of pooling data for post-treatment period with and without covariance analysis (n=252, period: 1961–80)

Gap(i)	Earlier period(p)	Later period (q) 1	2	3	4	5
—	0	100	148	155	160	163
0	1	121	262	280	314	333
	2	176	482	538	657	720
	3	179	513	606	718	789
	4	192	569	697	824	903
	5	200	574	714	827	946
1	1	202	331	367	382	389
	2	219	531	631	718	738
	3	224	575	705	790	820
	4	227	604	741	844	913
	5	229	605	746	884	917
2	1	143	273	306	327	346
	2	187	436	555	607	656
	3	197	488	597	646	729
	4	209	527	652	740	796
	5	203	524	669	724	814
3	1	147	266	295	308	321
	2	179	436	490	537	585
	3	184	449	515	594	605
	4	169	455	550	600	644
	5	164	468	548	615	696
4	1	157	285	318	340	368
	2	200	439	516	582	598
	3	197	455	559	590	642
	4	202	500	597	657	703
	5	211	507	615	649	715
5	1	160	287	324	359	364
	2	183	414	496	530	573
	3	183	455	525	568	615
	4	201	477	575	623	669
	5	209	491	567	620	659

times higher with five years' pre-treatment data, when data for five years in the post-treatment period was considered. From the above analysis, it may be concluded that, as and when the experiment progresses, it is desirable to use the progressive average yields for analysis, instead of analysing the annual yield data every year. That is to say, in the first year of the experiment, the first year's data can be analysed, while in the second year, the average yield for the first two years, and in the third year, the average yield for the first three years and so on, may be considered for analysis. Except in places where past yield records are regularly maintained, it is enough to consider only two years' data immediately prior to the commencement of the experiment, for covariance analysis. Present data did not support Govinda Iyer's (1957) suggestion of using three years' pre-experimental data for covariance analysis. Sen and Biswas (1966) have shown in tea, that adjustment of yield due to co-variance analysis generally ceased to be efficient after the first five years' experimentation in manurial trials. In the case of coconut, no such trend was noticed, though marginal decline in efficiency was visible, as the gap between pre-and post-experimental periods was increasing.

Acknowledgement

The authors are thankful to all those responsible for the collection of yield data from time to time and to their colleagues in the Division of Statistics for the computational help received.

References

ABEYWARDENA, V. 1962. Studies on biennial bearing tendency in coconut- The measurement of bienniality in coconut. *Ceylon Coconut Q.*, **13**: 112-115.

ABEYWARDENA, V. 1970. The efficiency of pre-experimental yield in the calibration of coconut yields. *Ceylon Coconut Q.*,**21**: 85-91.

ABRAHAM, T. P. and KULKARNI, G. A. 1963. Investigations on the optimum pre-experimental period in field experimentation in perennial crops. *J. Indian Soc. Agric. Stat.*, **15**: 175-183.

BUTTERS, B. 1964. Some practical considerations in the conduct of field trials with *robusta* coffee. *J. hort. Sci.*, **39**: 24-33.

CHAPAS, L. C. 1961. Plot size and reduction of variability in oil palm experiments. *Emp. J. exp. Agric.*, **29**: 212-224.

CUNNINGHAM, P. K. and BURRIDGE, R. C. 1957. The effect of yield variability of Amelonado cacao on the design of fertilizer field experiments in Ghana. *J. hort. Sci.*, **34**: 229-237.

EISENHART, C. 1947. The assumption underlying the analysis of variance. *Biometrics*, **3**: 1-21.

GOVINDA IYER, T. A. 1957. Statistical analysis of experimental yield data from coconut trees. *Indian Coconut J.*, **11**: 106-124.

LOTODE, R. and MULLER, R. A. 1974. Problems of experimentation with cocoa trees. In: *Phytophthora Diseases of Cocoa*. Gregory P.H. (ed.) Longman Group Ltd., London, 348 p.

PANKAJAKSHAN, A. S. 1960. A note on the relative contribution of genetic and environmental factors on the yield of uniformly treated coconut trees. *Indian Coconut J.*, **14**: 37-43.

SEN, A. R. and BISWAS, A. K. 1966. Some techniques of experimentation with tea in north-east India. *Expt. Agric.* **2**: 89-100.

SHRIKHANDE, V. J. 1957. Some considerations in designing experiments on coconut trees. *J. Indian Soc. agric. Stat.*, **9**: 82-99.

DISCUSSION

A. I. JOSE: Did you try the transformation $\sqrt{x+1}$? Can the $\sqrt{x+10}$ transformation be applicable to all types of experiments in coconut?

JACOB MATHEW: Yes, we tried $\sqrt{x+1}$ transformation along with transformations like $\sqrt{x+2/3}$, $\sqrt{x+20}$, $\sqrt{x+30}$, $\sqrt{x+40}$, $\sqrt{x+50}$, $\text{Log}_e(X+10)$, etc. $\sqrt{x+10}$ is expected to be applicable for all types of experiments irrespective of the type of treatment.

A. I. J.: What should be the minimum period for taking cumulative yield? Is it not dependent on the type of treatment?

J. M.: This aspect was not studied.

N. R. SHANTHAMALLAIAH. Can we take five years average yield data of coconut and subject it to analysis after $\sqrt{x+10}$ transformation? This I want to know because coconut needs $3\frac{1}{2}$ years from flower primordium to harvest of ripened nuts.

J. M. Our suggestion is to use the progressive averages for analysis. As and when the experiment progresses, we can go on adding the data. If you expect delay in getting response to treatments, we can leave a gap of one or two years, immediately after the commencement of the experiment.