SAMPLE REQUIREMENTS FOR ESTIMATING FRUITING AND YIELD OF COTTON

D. A. Wolfenbarger

Subtropical Crop Insects Research Unit, USDA, ARS,
509 West Fourth, Weslaco, TX 78596

ABSTRACT

For each of three sampling methods, the numbers of whole plants that had to be sampled to achieve statistically reliable estimates of populations of flower buds (squares), flowers, and bolls on cotton, were calculated using the equation n = (Coefficient of Variation/30%)^2. In addition, distributions of numbers of flower buds, flowers, and bolls were determined throughout the season and as a seasonal average, i.e., percentage of plants having such forms that fall with the discrete classes of 0, 1, 2, 3, etc. The sampling methods were used within a 0.4-ha plot; the "X" required sampling across rows diagonally from corner to corner of the plot; the "grid" required sampling down rows; and the "area" required the sampling of consecutive plants in a designated section in each quarter of the plot.

In another experiment, the number of plants of two cotton cultivars that had to be sampled to achieve reliable estimates of fruiting forms were calculated using a confidence interval equal to a fixed positive number.

The results indicated that sample sizes of about 15 to 100 plants/sampling date per area, regardless of the methods of sampling or calculating, or of the cultivar used, were adequate to achieve the desired accuracy.

INTRODUCTION

To estimate the fruiting characteristics and distribution of fruiting forms of a cotton crop, it is necessary to sample an adequate number of whole plants for flower buds (squares), flowers, or bolls in a field or defined area throughout most of the growing season. However, information on methods by which to determine the adequacy of sample sizes is limited.

Sterling and Pieters (1974) presented a sequential sampling plan for determining the extent of the need to sample for pinhead squares of cotton. These authors also showed that the sampling of more plants than was necessary to achieve statistically reliable estimates of populations of fruiting forms frequently occurred.

Dupnik et al. (1973) found no significant differences in mean numbers of fruiting forms when 22-100 cotton plants were sampled at various times and locations within 0.4-ha plots during a growing season, but the required sample sizes varied with the density of the fruiting forms.

Ideally, to test the adequacy of sample sizes for a given sampling technique, the actual number of fruiting forms in a given area should be known, thereby making it possible to calculate the accuracy or an estimate for each "subarea" or parcel sampled. Dupnik et al. (1973) and Wolfenbarger et al. (1980) applied this ideal measure by counting all plants in a 0.4 ha. Then they determined that statistically reliable estimates of the total number of plants were made on only one of nine sampling dates. Thus, if it is this difficult to obtain reliable results when sampling for objects such as stationary plants, it is reasonable to believe that the sampling of dynamic fruiting populations would be more difficult.

To acquire more information on this subject, the data on populations of flower buds, flowers, green bolls, and open bolls of two cultivars (Dupnik et al. 1973 and Wolfenbarger et al. 1970) were analyzed.

METHODS

The data of Dupnik et al. (1973), used here to estimate the number of plants to be sampled in 0.4-ha plots, were obtained by three methods of sampling: the "X", requiring sampling across rows diagonally from corner-to-corner of the square plot; the "grid", requiring sampling down ten rows within the plot; and the "area", requiring the sampling of consecutive plants in an area near the center of the plot. The data for flower buds, flowers, and bolls on individual plants were analyzed using the equation $n = \text{Coefficient of Variation}/30\%)^2$ (Harcourt 1961, Mukerji and Harcourt 1970, Wolfenbarger et al. 1980).

Frequency distributions of fruiting forms that occurred on five sampling dates over a 2-month period during the growing season were determined, and X^2 analysis was used to test for the presence of skewness and kurtosis (Snedecor 1946). When the values for these indices were less than 2.57 at the 1% level of probability, the distribution was Gaussian. The exponential regression equation $Y = ae^{-bX}$ was used to determine the class frequencies of percentages of plants that bore 1, 2, 3...10 flower buds, flowers, or bolls on each sampling date.

RESULTS AND DISCUSSION

On Julian date 146, 25% of the plants had four flower buds. Ten flower buds were present per plant on 23% and 15% of the plants on Julian 173 and 184, respectively. Flower bud populations on Julian dates 146, 173, and 184 formed normal (Gaussian) distributions (Fig. 1), whereas on Julian dates 209 and 210, when the plants were mature and more than 50% had fewer than four flower buds each, the distributions were not normal.

The mean number of flower buds and number of plants that had to be sampled to obtain reliable estimates were less when Gaussian, rather than curvilinear, distributions existed (Fig. 1, Table 1), thus indicating a relationship between

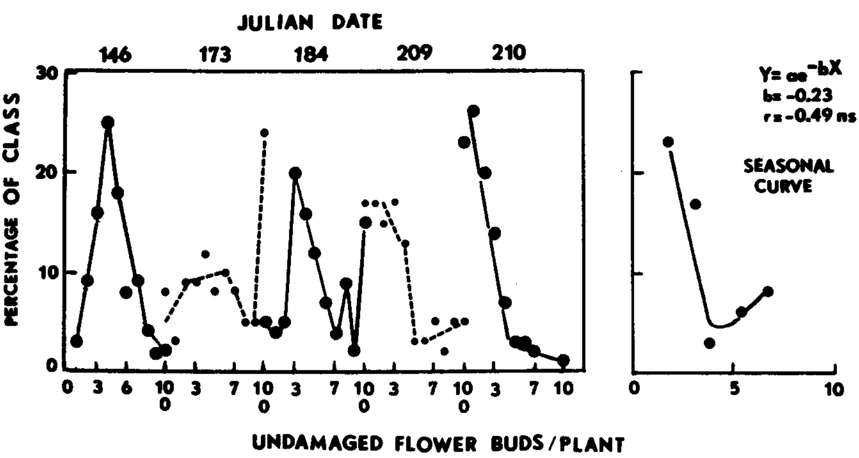

FIG. 1. Frequency distribution of flower buds undamaged by insect feeding per whole plant on 5 dates using the "area" method of sampling (left). Exponential regression of seasonal curve, slope (b) and correlation coefficient (r); ns = not significant (right).

TABLE 1. Sampling Requirements for Three Sampling Methods Based on the Mean Number of Fruiting Forms of Cotton as Determined by n (No. of Plants to be Sampled) = $(CV/30\%)^2$.[a]

	Number of plants to be sampled					
	Area		Grid		X	
Julian date	Mean no./plant	Range of n	Mean no./plant	Range of n	Mean no./plant	Range of n
			Flower buds/plant			
122	0.5	67-189	0.4	31[b]	0.25	47-120
138	0.11	8-27	1.0	15[b]	2.5	4-5
146	2.5	7-8	4.0	4-5	4.0	4-4
173	5.5	7-7	4.0	3-4	7.0	5-7
184	7.5	6-7	7.5	5-9	7.5	4-6
209-210	3.5	16-20	2.5	11-18	3.0	10-14
			Flowers/plant			
146	0.15	67-135	0.03	556[b]	0.03	109-578
173	0.43	22-36	0.35	26-35	0.4	37-39
184	0.31	31-44	0.2	81-155	0.2	57-65
209-210	0.31	22-56	0.35	39-56	0.45	18-25
			Green bolls/plant			
146	0.55	90-867	0.09	93-131	0.065	109-283
173	1.5	11-20	1.5	9-15	1.5	14-18
184	2.5	8-10	2.5	8-9	2.0	8-12
209-210	4.5	10-13	3.0	5-8	3.5	4-6
			Open bolls/plant			
146			0.09	93-131	0.065	109-283
173			1.5	9-15	1.5	14-18
184			2.5	8-9	2.0	8-12
209-210			3.0	5-8	3.5	4-6

[a] Method of Harcourt (1961) and Murkerji and Harcourt (1970).
[b] Data from sampling in one replicate.

the sample size estimate and the type of distribution. The curves displayed on Julian dates 146 and 184 (excluding the last data point of 10 undamaged flower buds per plant) were not skewed at the 5% level of probability. On the respective dates, the X^2 values for skewness were 0.93 and 0.43, and for kurtosis, 0.12 and 1.32. However, on Julian date 173 the X^2 values for skewness and kurtosis were 2.3 and 6.9, respectively.

When the regression $Y = ae^{-bX}$ was applied to the seasonal totals of classes of flower buds, the slope was flat and the correlation coefficient (r) was -0.49 which was not significant at the 5% level of probability (Fig. 1). Hence flower buds per sample displayed a Gaussian distribution until reduced numbers at the end of the season caused them to show an exponential distribution.

Populations of flowers ranged from 0-4 per plant, and normal distributions did not occur on any sampling date (Fig. 2). When the regression $Y = ae^{-bX}$ was applied to the seasonal totals, the data produced a correlation coefficient of -0.97, significant at the 5% level of probability, and a slope of -1.04 (Fig. 2).

Bolls were found to be in Gaussian distribution on Julian dates 184, 209, and 210 (Fig. 3). Values for skewness and kurtosis on the same dates were 0.27 and 1.6, 0.7 and 0.4, and 0.04 and 1.8, respectively.

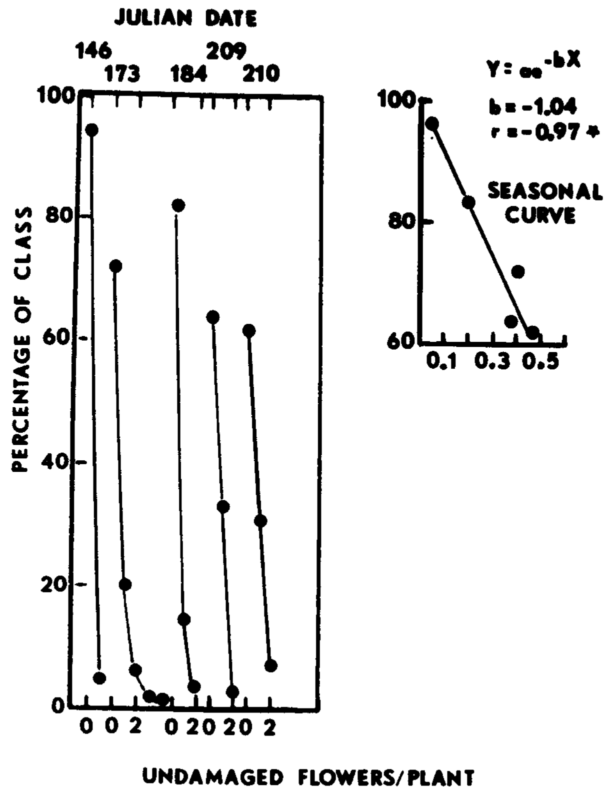

FIG. 2. Frequency distribution of flowers undamaged by insect feeding per whole plant on 5 sampling dates using the "area" method of sampling (left). Exponential regression of seasonal curve, slope (b) and correlation coefficient (r); * = significant at the 1% level of probability (right).

The equation $Y = ae^{-bX}$ fit the data on the seasonal totals of bolls per plant (Fig. 3); the "r" was significant at the 5% level of probability, and the slope approached -1.0, a value similar to that shown for flowers. The distribution was exponential on Julian dates 146 and 173 when the mean numbers of bolls per plant were very low.

When these distributions were used to estimate sample sizes, it was found that the greater the mean of the object sampled, the lower the number of samples necessary (Table 1). Also, the mean numbers of flower buds, flowers, and green bolls per plant obtained by the three sampling methods were similar.

Using a confidence interval equal to a fixed positive number (Karandinos 1976), estimates of the numbers of plants to be sampled were made from data collected in 1978 on numbers of pinhead, 1/3-grown, and full-grown flower buds, and small, medium, and large bolls on plants of a short-season (Tamcot SP-37) and a standard-season (Stoneville ST-213) cultivar (Wolfenbarger et al. 1979). A fixed positive number of 1 was used for the confidence interval. The standard deviations, as percent of the mean of each index, were equal to or greater than their means on about 50% of the sampling dates. Also, by this method of calculation, no more than 72, 79, or 121 plants of either cultivar had to be sampled to obtain reliable estimates of populations of pinhead, 1/3-grown, and full-grown flower buds, respectively (Table 2). On the other hand, reliable estimates for sample sizes of small, medium, and large bolls could be obtained from samplings of 34, 31, and 36 plants, respectively. No more than 51 plants had to be sampled to obtain adequate estimates of open bolls.

Hence, regardless of the method of calculation employed, the manner in which sampling was conducted, or the cultivar used, roughly 15 to 100 whole plants (depending largely on the fruitload at the time of sampling) were sufficient to provide statistically reliable estimates of the fruiting forms of the cotton.

TABLE 2. Sampling Requirements for Fruiting Forms of Cotton Based on the Mean and Percent Standard Deviation and a Fixed Confidence Interval[a] ("Area" Method of Sampling).

	Cultivar				No. of plants to be sampled	
	SP-37		ST-213			
Julian date	Mean no.	s as % of mean	Mean no.	s as % of mean	SP-37	ST-213
	Pinhead flower buds/plant					
125	2.2	55	1.5	53	6	2
130	3.7	51	3.2	44	14	8
137	5.1	39	4.9	43	16	16
144	5.3	51	6.6	58	28	55
152	7.0	41	8.2	51	30	66
160	4.3	100	3.2	103	72	42
165	1.6	218	1.5	220	48	41
172	0.6	300	1.3	169	13	18
179			0.9	354		46
187	0.08	500	0.1	300	1	1
193			0.05	400		1
206			0.1	500		1
	1/3 grown flower buds/plant					
137	5.1	34	5.0	52	16	26
144	5.9	44	6.2	52	26	40
152	6.2	66	6.1	74	65	79
160	2.5	112	3.4	100	30	43
165	1.3	192	1.8	156	25	30
172	0.6	300	0.8	238	13	14
179			0.6	350		18
206	0.04	500	0.1	200	1	1
	Full grown flower buds/plant					
137	0.7	129	0.9	189	3	11
144	0.7	729	5.1	63	100	40
152	6.1	92	5.6	59	121	43
160	3.6	97	5.8	79	46	81
165	2.3	213	3.5	100	92	47
172	1.0	250	1.5	140	23	16
179			0.3	333		4
193			0.2	350		2
	Small bolls/plant					
144	0.2	250	0.2	200	1	1
152	2.7	30	2.0	65	15	7
160	4.1	59	3.0	80	23	22

(continued)

TABLE 2. (Continued.)

Julian date	Cultivar SP-37 Mean no.	SP-37 s as % of mean	ST-213 Mean no.	ST-213 s as % of mean	No. of plants to be sampled SP-37	ST-213
			Small bolls/plant (continued)			
165	2.6	104	4.0	75	28	34
172	1.7	124	1.6	125	18	15
179	1.1	155	1.5	140	1	17
187	0.2	250	0.2	250	1	1
200	0.04	500	0.05	400	1	1
			Medium bolls/plant			
152	0.7	129	0.2	250	3	1
160	1.7	83	2.0	80	8	10
165	2.8	71	3.1	71	16	18
172	3.6	69	3.7	76	24	31
179	2.7	56	2.7	56	9	9
187	1.5	67	1.6	100	4	10
193	0.6	167	0.3	200	4	2
200	0.3	267	0.1	300	3	1
206	0.1	400	0.1	300	1	1
			Large bolls/plant			
160	0.7	128	1.1	109	3	6
165	0.7	171	0.6	217	6	6
172	0.6	200	1.0	140	6	8
179	2.6	81	2.8	111	17	36
187	3.6	69	4.5	47	23	18
193	1.8	111	3.1	94	15	33
200	0.5	260	1.2	175	6	17
206	0.1	400	0.3	300	1	3
			Open bolls/plant			
187	1.3	85	0.6	150	5	3
193	3.5	66	2.9	86	21	23
200	5.1	45	4.9	47	21	20
206	5.8	41	5.8	62	22	51

a/ Method of Karandinos (1976).

ACKNOWLEDGMENT

The author thanks Dr. Del Val Petersen, Biometrician, USDA, ARS, for preparing an analytical program for the programmable calculator to determine sample sizes.

LITERATURE CITED

Dupnik, T. D., D. A. Wolfenbarger, J. A. Harding, and T. W. Fuchs. 1973. Sample method and size comparisons for fruit and insects in cotton. Texas Agric. Expt. Stn. Misc. Pub. 1121, 18 pp.

Harcourt, D. G. 1961. Design of a sampling plan for studies on the population dynamics of the diamondback moth, Plutella maculipennis (Cort.) (Lepidoptera: Plutellida). Can. Entomol. 93:820-30.
Karandinos, M. G. 1976. Optimum sample size and comments on some published formulae. Bull. Entomol. Soc. Amer. 22:417-21.
Mukerji, M. K. and D. G. Harcourt. 1970. Design of a sampling plan for studies on the population dynamics of the cabbage maggot, Hylemya brassicae (Diptera: Anthomyiidae). Can. Entomol. 102:1513-18.
Snedecor, G. W. 1946. Statistical Methods. 4th Edition. Iowa State Univ. Press. 485 pp.
Sterling, Winfield L., and E. P. Pieters. 1974. A sequential sampling package for key cotton arthropods. Dept. Entomol., Texas A&M Univ. Tech. Rep. No. 74-32, 28 pp.
Wolfenbarger, D. A., R. H. Dilday, and J. W. Davis. 1979. Fruiting requirements of short-season cottons to avoid damage by the boll weevil in the Lower Rio Grande Valley. Proc. Beltwide Cotton Res. Conf. In Cotton Improv. Conf., Phoenix, AZ, pp. 88-90.
Wolfenbarger, D. A., H. Del Var Petersen, T. D. Dupnik, and J. A. Harding. 1980. Cotton: Sample sizes required to estimate number of plants and fruiting forms by 3 systems of sampling. Southwest. Entomol. 5:93-8.

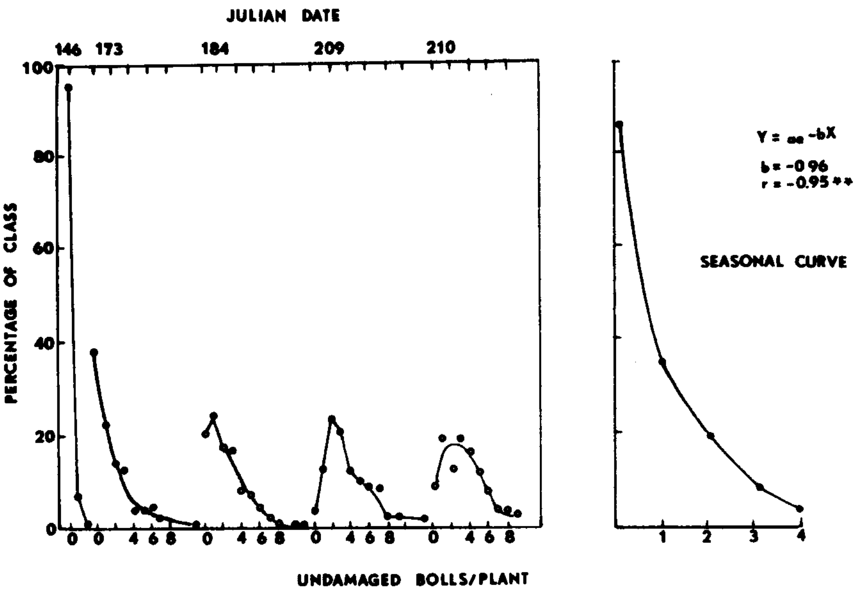

FIG. 3. Frequency distribution of bolls undamaged by insect feeding per whole plant on 5 dates using the "area" method of sampling (left). Exponential regression of seasonal curve; slope (b) and correlation coefficient (r); * = significant at the 1% level of probability.