THE SEQUENTIAL SAMPLING OF *PSOROPHORA COLUMBIAE* LARVAE IN RICE FIELDS

R. E. McLAUGHLIN[1], M. A. BROWN[2] AND M. F. VIDRINE[3]

U. S. Department of Agriculture, Agricultural Research Service, Insects Affecting Man and Animals Research Laboratory, P.O. Box 14565, Gainesville, FL 32604

ABSTRACT. Sequential sampling plans were developed for larval populations of *Psorophora columbiae* developing in rice fields during the growth of rice for a 2nd crop in southwestern Louisiana. Two types of plans were developed: 1) sequential counting plans to estimate the true mean of a population with a prespecified coefficient of variation and 2) sequential decision plans to assist in deciding whether the mean larval count/dip lies outside preselected upper or lower levels. The procedure by which the plans were developed consisted of: 1) conventional larval sampling by dipping along rice field levees that divided each field into pans; 2) counting the number of 2nd through 4th instar larvae observed in two dips taken at each sampling location; and 3) determination of the appropriate statistical parameters from which probability curves, number of samples required, and cumulative larval totals for specific sampling plans could be derived. Both sequential sampling plans are based upon distributions attendant to naturally occurring populations.

INTRODUCTION

The design of a valid and practical sampling procedure is based upon the principle of specified precision at minimum cost (Cochran 1977). Fixed sample size plans are efficient only within a narrow portion of a broad population size range, usually resulting in an inordinately excessive number of samples when the population density is very low and too few samples being taken when populations are very dense. Thus, the desired level of precision often is either grossly exceeded or not attained. More importantly, the precision of the estimate of variance is seldom estimated for a particular group of samples and thus the variance of estimated sample size required is ignored.

Development of valid sampling plans for mosquito larvae must be based upon probability theory and the spatial distribution of larvae. Determination of the fit of the data to a particular type of spatial distribution then allows development of sequential sampling plans which provide optimum levels of precision for classification of the population into density categories.

The lack of data relating mosquito larval density to subsequent adult density makes absolute larval estimates of less importance to control agencies than the knowledge that the true density has a high probability of being either greater or less than some preselected levels, separated by some range. The rationale is that densities greater than a certain upper threshold level are clearly of concern, but that densities below a certain level probably do not justify further sampling; densities lying between the two levels may require further sampling. Such an approach to population surveillance might provide increased efficiency in the process of classifying populations into preselected density categories and of assessing the precision of that classification.

An excellent presentation of the various types of sequential decision plans for the entomologist who wishes to be spared detailed mathematical treatises was published by Onsager (1976). Sequential sampling plans for mosquito larvae have been published by Mackey and Hoy (1978) and Stewart et al. (1983), who suggested a negative binomial distribution for *Culex tarsalis* Coquillett in rice fields based on goodness of fit tests. The spatial distribution pattern of *Psorophora columbiae* (Dyar and Knab) larvae in rice fields was found to be consistently patchy and clumped, (Andis and Meek 1984), which is indicative of negative binomial distribution patterns. They further stated, based on their information regarding dispersion patterns of the larvae, that the development of more precise strategies for estimating rice field mosquito larval populations could now be developed.

The objectives of this study were: 1) to calculate parameters from which any number of sequential decision plans could be developed once upper and lower limits of population density and the desired levels of precision had been selected by a user; and 2) to calculate the parameters and derive a sequential counting plan for estimation of the mean, with predetermined coefficients of variation according to the method of Kuno (1969).

[1] Insects Affecting Man and Animals Research Laboratory, U.S. Department of Agriculture, Agricultural Research Service, P.O. Box 14565, Gainesville, FL 32604.

[2] South Central Family Farms Research, U.S. Department of Agriculture, Agricultural Research Service, Rt. 2, Box 144-A, Highway 23 South, Booneville, AR 72927.

[3] Division of Science, Louisiana State University at Eunice, Eunice, LA 70535.

MATERIALS AND METHODS

Sampling procedure. The sampling tool was a dipper of ca. 400 ml capacity, fitted with a handle ca. 1.2 m long. This was immersed in such a manner as to minimize escape of larvae alarmed by water ripples or shadows. The number of captured larvae and the stage of the majority of larvae were then recorded. Each sample consisted of the sum of the numbers of larvae in two dips taken at a specific point in a pan, which is a subdivision of the rice field created by earthen levees for water management. Distance between sampling locations was 70 steps. When pans were too small to permit at least 20 sampling locations in a pan, samples were collected at 35-step intervals. Samples were taken along and within 5 m of the levees around the entire periphery of each pan. A single line of samples was taken down the center for the length of the pan in pans less than ca. 35 steps wide. Pans were sampled 2 to 4 days after flooding. Data from a pan were not used if the majority of the insects were in the first larval instar or the pupal stage. First instar larvae tend to congregate around debris, resulting in highly clumped aggregates of very small larvae. The presence of pupae indicated that some adult emergence may have occurred, which would invalidate the counts. Therefore, 2nd through 4th instar larvae were used for determination of sequential sampling schemes.

Data base. Data were collected from rice fields that were reflooded after the first harvest to stimulate regrowth of the cut stubble for production of a second rice crop during late summer of 1982 and 1983. All of the fields were contained within an area of 117 km^2 around the town of Jennings, LA. Sixteen fields (236 ha) were selected at random as they were being flooded. All 221 pans in the 16 fields were sampled from July 30 through September 7, 1982. In 1983, 115 pans (2 to 4 per field) were sampled from 41 fields located in the same area to increase the number of fields in the data base. The mean count from each "pan within field" was chosen as the observational unit for the statistical analysis, because the statistical inference would be to a random pan in a random field throughout the time period of the study.

RESULTS AND DISCUSSION

Parameters of the negative binomial distribution (the mean, and the aggregation coefficient k) were obtained by the methods given by Wald (1945), Bliss and Owen (1958), and more recently by Allen et al. (1972). Provisional estimates of a common k for each pan were calculated, and heterogeneity of the k values was tested. As suggested by Bliss and Owen (1958), the reciprocal of 1/k was used in provisional estimates of a common k. Estimates of a common k from 2nd instar vs. older instar samples were compared and found to be statistically homogeneous; thus, data for instars 2 through 4 were consolidated in further development of the methods. Separate estimates of k were obtained for each year's data and tested for homogeneity; the hypothesis that the two estimates were not different could not be rejected (F value = 0.45; $P > F = 0.64$). Data from both years were then pooled to yield an estimate of 1.26 for the common k ($n = 332$, $P > F = 0.0001$, standard error of the estimate = 0.0595, coefficient of variation of 1/k = 0.07).

Sequential decision plans were then developed, including an operating characteristic (OC) curve and an average sample number (ASN) curve. Calculation of a specific sequential decision plan was accomplished by following the procedures of Wald (1945). Two means representing the lower ($\bar{X}_1$) and upper ($\bar{X}_2$) limits were selected for calculation of any specific plan. Values that dictate rejection of the hypothesis that the true number of larvae $> \bar{X}_1$ (H_1) or $< \bar{X}_2$ (H_2) are given by $\hat{Y}_1 = a_1 + bx_1$, for H_1, and $\hat{Y}_2 = a_2 + bx_2$, for H_2, where: $\hat{Y}_i$ = total larvae needed for acceptance of H_i (i = 1,2); $a_1 = -\{[\log[(1 - \alpha)/\beta]]/[\log[(P_2Q_1)/(P_1P_2)]]\}$; $a_2 = \{\log[(1 - \beta)/\alpha]/\log[(P_2Q_1)/(P_1Q_2)]\}$; $b = k\{[\log(Q_2/Q_1)]/[\log(P_2Q_1/P_1Q_2)]\}$; α = desired probability of incorrect acceptance of H_1; β = desired probability of incorrect acceptance of H_2; $\bar{x}_i$ = critical minimum ($\bar{x}_1$) or maximum ($\bar{x}_2$) larval count (i = 1,2); $P_i = \bar{x}_i/k$, (i = 1,2); $Q_i = 1 + P_i$, (i = 1,2); k = aggregation coefficient estimated from data; and x = total number of two-dip samples examined.

According to the established procedures, the maximum ASN required for a decision will be close to $-\{[a_1a_2]/[(b^2/k) + b]\}$, and that ASN will be required when the true density of larvae is approximately equal to b.

The OC curve gives the probabilities of accepting the hypotheses as a function of the true larval count (x_1) and is normally given for H_1, (H_2 equals 1-H_1). Four points can be easily calculated for this curve:

x_1	Probability
0	1
$\bar{x}_1$	$1 - \alpha$
b	$a_2/(a_2 - a_1)$
$\bar{x}_2$	β

Sequential counting plans were developed following the method described by Allen et al. (1972), which provides an estimate of the pop-

ulation mean with a fixed coefficient of variation. Reiteration of the referenced articles here will assist the reader to more readily understand the derivation of specific counting plans after selection of a desired level of precision of the estimate of the true larval density in a pan. Let: m = an estimate of the population mean (= T_n/n), T_n = the total number of larvae in n samples, S_m = the standard error of m (= $\sqrt{S^2/n}$), S^2 = the estimate of the population variance, n = the number of samples taken, and C = the coefficient of variation of the mean (S_m/n).

Thus,

$$C = nS^2/(T_n)^2. \quad [1].$$

Equation [1] is solved for T_n to determine the total number of larvae to be collected as a function of sample size n. The sample variance, which changes in value more rapidly than the mean in spatially aggregated distributions, in the numerator of equation [1] presents a problem. Allen et al. (1972) substantiated the proposals by Anscombe (1949) and Kuno (1969), that the sample variance be expressed as a function of the mean so that $S^2 = f(m)$, taking the form of $S^2 = (\alpha + 1)m + (\beta - 1)m^2$ where α and β are constants. Assuming a linear relationship between mean crowding and the mean, as defined by Lloyd (1967), the practical solution is attained by obtaining least squares fit to the quadratic $S^2 = a_1m + a_2m^2$, where a_1 and a_2 are constants. With replacement of S^2 by this quadratic, equation [1] then assumes the form: $C = \sqrt{\{[n(a_1T_n/n) + (a_2T_n^2)/n^2]/T_n^2\}}$. Solving for T_n gives:

$$T_n = a_1/C^2 - (a_2/n) \quad [2].$$

A curve analogous to the ASN function of the sequential decision plan can be derived by dividing equation [2] by n and solving for n,

$$n = (a_1/mC^2) + (a_2/C^2) \quad [3].$$

Such a solution is used to examine the effect of different coefficients of variation (C) on the expected number of samples to be taken. Equation [3] is thus useful in selection of a C value while taking into account the sampling resources available, past experience with the expected C values and the experimental design of the test.

The computation of a sequential sampling plan involves selection of the upper and lower levels of larval density and an appropriate level of precision. A specific sequential sampling plan can then be easily computed from the formulae given above, using the values of k, $\bar{x}1$, $\bar{x}2$, α and β. The curve for any specific plan shows the number of samples that is expected to yield a decision or classification. Therefore, one can consider the labor and other resources available for sample collection along with the factors of larval density and precision. The ASN for a variety of combinations of the upper and lower levels of larval density and levels of precision are shown in Table 1. Additional ASN values can be computed from earlier formulae if one desires to choose other upper and lower larval density levels.

Poston et al. (1983) discussed the question of economic injury levels and action thresholds for treatment in relation to four categories of pest management decisions representing four degrees of development of reliability or levels of specialization. The least developed and tested one was termed the nominal threshold and was recommended as a starting point for the development of more elegant thresholds. They defined a nominal threshold as one based on the experience of the pest manager. In the case of *Ps. columbiae*, nominal threshold levels for a sequential decision plan were chosen from our 4 years of experience to assist in the selection of fields to be used in larvicidal tests of *Bacillus thuringiensis* serotype H-14 (McLaughlin and Vidrine 1984). These levels will be used here to illustrate the practical employment of sequential decision plans.

Lower and upper limits of 1 and 3 larvae per sample, respectively, represented to us a fair compromise between time and accuracy. Our experience showed that most of the fields with more than 1.5 larvae/dip (3 per sample) usually had considerably higher larval densities, and fields with less than 0.5 larvae/dip (1 per sample) seldom had high larval counts in any pan.

Reference to Table 1 shows that the ASN for selected precision levels of 0.05, 0.10 or 0.25

Table 1. Sequential decision plans for several precision levels of α & β, and lower ($\bar{X}_1$) and upper ($\bar{X}_2$) limits of the mean number of *Psorophora columbiae* larvae per sample (2 dips/sample), giving the associated ASN (maximum number of samples to be taken) for each combination.

Plan no.	Level of α, β	Mean no. larvae/sample $\bar{X}_1$ Lower	$\bar{X}_2$ Upper	ASN
1	0.05	1.00	1.50	85
2	0.05	1.00	2.00	27
3	0.05	1.00	3.00	10
4	0.10	0.50	1.50	8
5	0.10	1.00	1.50	47
6	0.10	1.00	2.00	15
7	0.10	1.00	3.00	6
8	0.10	2.00	3.00	35
9	0.25	0.50	1.50	2
10	0.25	1.00	1.50	12
11	0.25	1.00	2.00	4
12	0.25	1.00	3.00	2
13	0.25	2.00	3.00	9

would be 10, 6 or 2, respectively. The values for the cumulative number of larvae required as sampling progressed in order to arrive at a decision and cease sampling are presented in Table 2 as an example of the use of the plan selected with a precision of 0.05. In usage, one enters the field and commences to sample, entering the total number of larvae in the first sample of two dips on the first line opposite sample no. 1. If the larval count is 8 or more, sampling could be discontinued in that pan. Sampling may be terminated at any time the cumulative larval count exceeds the upper limit number. If the count was less than 8, another sample would be taken further along in the pan. The number of larvae in that sample would then be added to the previous count and entered on the 2nd line. Sampling could now be terminated if the count was 10 or more larvae. One can see that the fourth sample must be taken before any decision can be made as to the true mean being less than 1 larva per sample. At any time after the third sample that the cumulative larval total is less than or equal to the lower number sampling required can be terminated in that pan with the confidence that the true mean count is less than 1 larva per sample with the specified precision. Sampling continues until the maximum number of samples has been taken for that plan if the cumulative number of larvae never falls below the lower limit or exceeds the upper limit. At that point in the sampling procedure the statement can be made that no decision can be made as to whether the true mean lies outside the selected range.

This procedure has the advantage of permitting the assessment of the population with the optimum expenditure of sampling effort and with the knowledge of the probability of making the incorrect decision when the true mean is of less importance than the determination that it is less than or greater than preselected levels of interest. On the other hand, experimental testing of larvicides or assessment of other events directly influencing larval densities require more precise evaluation and an assessment of the true population mean is required for comparisons.

Table 2. Sequential decision plan for field use giving cumulative *Psorophora columbiae* larvae counts per sample (2 dips per sample; α and $\beta = 0.05$). Lower and upper limits of the mean number of larvae per sample = 1.0 and 3.0. Cumulative counts outside the limits indicate that the true mean lies either below (lower limit) or above (upper limit) the selected number. Missing value represents an impossible solution ($T_n < 0$).

	Cumulative number of larvae in 2-dip sample		
Sample no.	For lower limit of <1.0/sample	In 2-dip sample	For upper limit of >3.0/sample
1	—	—	8
2	—	—	10
3	—	—	11
4	1	—	13
5	2	—	15
6	4	—	14
7	6	—	18
8	7	—	20
9	9	—	22
10	11	—	24

Table 3. Sequential counting plan for *Psorophora columbiae* larvae (2nd-4th instar, sum of larvae in two dippers per sample); for determining the estimate of the true mean larval density with associated coefficient of variation (C-value).

Larval total	Sample no.	Cumulative number of larvae per sample needed to achieve C-value =							
		0.75	0.60	0.50	0.40	0.30	0.25	0.20	0.10
—	1	59	309	—	—	—	—	—	—
—	2	37	73	158	—	—	—	—	—
—	3	32	58	102	265	—	—	—	—
—	4	31	53	86	181	—	—	—	—
—	5	30	50	79	152	538	—	—	—
—	6	28	48	75	137	391	—	—	—
—	7	29	47	72	129	327	828	—	—
—	8	28	46	70	123	291	632	—	—
—	9	28	46	69	119	268	534	—	—
—	10	28	45	68	115	253	475	—	—
—	15	27	44	65	107	215	356	774	—
—	20	27	43	63	103	200	317	609	—
—	25	27	43	62	101	192	297	540	—
—	30	27	43	62	99	187	287	502	—
—	35	27	42	62	98	183	277	478	13,373
—	40	27	42	61	98	181	272	461	6,750
—	45	27	42	61	97	179	267	449	4,843
—	50	27	42	61	97	177	264	440	3,950

Sequential counting plans provide the optimum method for those conditions.

Regression analysis from the mean-variance pairs for subsequent calculation of sequential counting plans resulted in a value of $a_1 = 14.8541$, $a_2 = 0.31197$, with an intercept of 2.2686 larvae/dip (4.5372 larvae/sample). The means ranged from 0.0588 to 28.810 per dip. Solution of equation [2] then produced the total number of larvae required in n samples to arrive at the desired coefficient of variation of the mean (C-value). Several coefficients are tabulated in Table 3 as a field-use form. The stop-sampling-lines, equation 2, are illustrated in Fig. 1 for various C-values.

Examination of the stop-sampling-lines in Fig. 1 reveals that the higher the mean larval density the greater will be the precision of the estimate of the mean. The stop-sample lines descend sharply during the initial few counts and very quickly parallel the x-axis. Highly accurate estimates of the mean of very low population densities may be beyond manpower resources normally available for sampling during experimental tests. A practical approach is suggested by the general flattening of the curves by

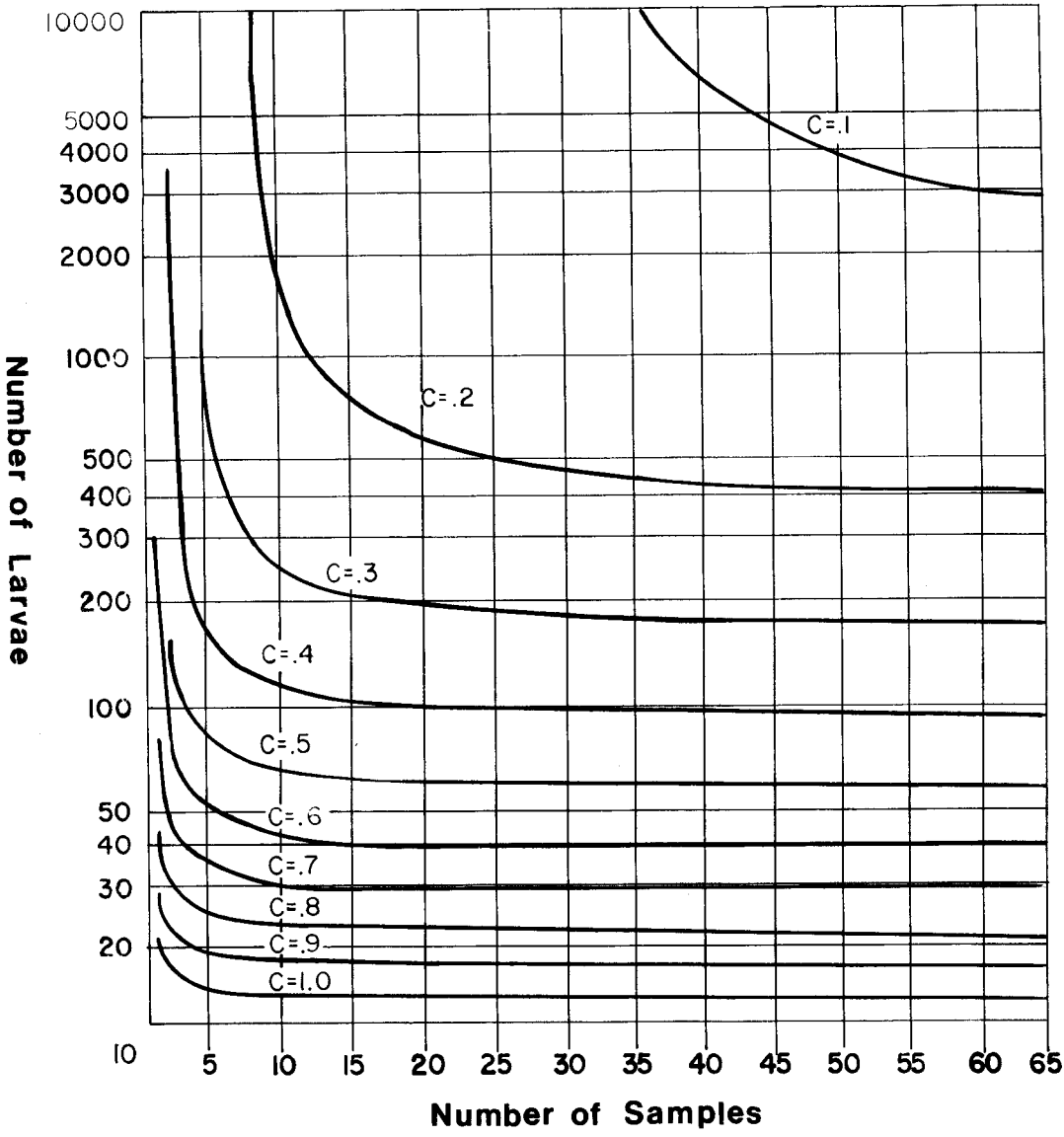

Fig. 1. Stop-sampling lines for several coefficients of variation of the mean number of larvae as a function of the number of 2-dip samples (x axis) and the total number of larvae (y axis).

the 10th sample for C-values of 0.4 or higher. C-values of 0.2 to 0.3 have a flattening of the curve at about 20 samples. This means that greatly increasing the number of samples does not result in parallel increases in sampling precision. In other words, decreasing efficiency of sampling resource expenditure occurs as sample size increases beyond these numbers.

The use of a sequential counting plan is illustrated in the following example. Let us suppose that a test has been planned to assess the change in the mean larval density of *Ps. columbiae* larvae due to some event. The assumptions required for validity of a sequential counting plan require that the level of precision be selected prior to actual sampling. The values of the cumulative number of larvae required to achieve an estimate of the true mean with various levels of precision are presented in Table 3. Examination of the values indicates that the larval population must be relatively high in order for us to achieve the stop-sampling-line within any reasonable number of samples. Prior experience must guide selection of the C-value by any user, but for purposes of example, we shall choose that of 0.6. The field is entered and the first sample of 2 dips is taken and the total number of 2nd-4th instar larvae recorded opposite sample no. 1 on the worksheet (Table 3). Let us assume sample counts of 2, 4, 7, 0, 5, 9, 6, 11, 0 and 4. The cumulative total number of larvae would be 4, 6, 13, 13, 18, 27, 33, 44, 44 and 48 for the 10 samples. The counting plan for the 0.6 C-value in Table 3 shows that sampling could be terminated at the 10th sample. The result would show a mean of 4.8 larvae per sample (2.4/dip) with a coefficient of variation of 0.6.

In the design of larval control experiments in which readily discernible differences between means are desired, common practice has often been to select pans with high larval densities. The higher densities result in achievement of a selected C-value with fewer samples than when very few larvae are present. However, by use of the sequential counting approach, an acceptably higher degree of precision can be achieved even with lower density plots by increasing sampling intensity. The additional sampling can be determined at the time of surveillance, and sampling intensity may differ between plots when using this approach as long as the prerequisite for prior selection of the C-value has been met. Thus the larval density required for acceptability of a pan for the experiment can be established during the design phase. This process is a useful application of sequential counting plan methodology.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of John Billodeaux, Director, Jefferson Davis Parish Mosquito Abatement District No. 1, Jennings, LA, and that of his staff, and the assistance of Colin Fake and the Graphic Arts and Drafting Department, Sowela Technical Institute, Lake Charles, LA, in the preparation of the figure.

REFERENCES CITED

Allen, J. D., D. Gonzalez and D. V. Gokhale. 1972. Sequential sampling plans for the bollworm, *Heliothis zea*. Environ. Entomol. 1:771–780.

Andis, M. D. and C. L. Meek. 1984. Bionomics of Louisiana riceland mosquito larvae. II. Spatial dispersion patterns. Mosq. News. 44:371–376.

Anscombe, F. J. 1949. The statistical analysis of insect counts based on the negative binomial distribution. Biometrics 5:165–173.

Bliss, C. I. and A. R. G. Owen. 1958. Negative binomial distributions with a common k. Biometrika 45: 37–58.

Cochran, W. G. 1977. Sampling techniques. John Wiley and Sons, New York. 427 pp.

Kuno, E. 1969. A new method of sequential sampling to obtain the population estimates with a fixed level of precision. Res. Popul. Ecol. 11:127–136.

Lloyd, M. 1967. Mean crowding. J. Anim. Ecol. 36: 1–30.

Mackey, B. E. and J. B. Hoy. 1978. *Culex tarsalis*: Sequential sampling as a means of estimating populations in California rice fields. J. Econ. Entomol. 71:329–334.

McLaughlin, R. E. and M. F. Vidrine. 1984. Comparison of the *Bacillus thuringiensis* serotype H-14 formulation against *Psorophora columbiae*. Mosq. News. 44:351–356.

Onsager, J. A. 1976. The rationale of sequential sampling with emphasis on its use in pest management. USDA Tech. Bull. 1526, 19 pp.

Poston, F. L., L. P. Pedigo and S. M. Welsh. 1983. Economic injury levels: reality and practicality. Bull. Entomol. Soc. Am. 29:49–53.

Stewart, R. J., C. H. Schaefer and T. Miura. 1983. Sampling *Culex tarsalis* (Diptera: Culicidae) immatures on rice fields treated with combinations of mosquitofish and *Bacillus thuringiensis* H-14 toxin. J. Econ. Entomol. 76:91–95.

Wald, A. 1945. Sequential tests of statistical hypotheses. Ann. Math. Stat. 16:117–186.