Variable-Intensity Sampling: A New Technique for Decision Making in Cabbage Pest Management[1]

C. W. HOY, C. JENNISON,[2] A. M. SHELTON, AND J. T. ANDALORO[3]

Department of Entomology, New York State Agricultural Experiment Station, Cornell University, Geneva, New York 14456

ABSTRACT

J. Econ. Entomol. 76: 139–143 (1983)

A variable-intensity sampling method was developed to efficiently monitor all cabbage pests in commercial fields in New York while concentrating on the most important pests, *Pieris rapae* (L.) and *Trichoplusia ni* (Hübner). The field is walked in a V-shaped pattern which is divided into 10 segments. Along each segment, one to four randomly selected cabbage plants are inspected, the number depending on an estimate of population density of *P. rapae* and *T. ni* derived from previous segments. Computer simulations of this procedure were run, using random numbers from the appropriate negative binomial distribution, and the 95% confidence intervals calculated contained the true mean with frequencies between 93 and 97%. In a commercial cabbage field, a comparison of estimates of the means obtained by variable intensity sampling and fixed sample sizes of 40 and 264 plants revealed only small differences.

Pest management should rely on some method of estimating pest population densities for treatment decisions. Sampling procedures that utilize a fixed sample size are inefficient when the pest population density is either very low or very high, because the sampling intensity is greater than necessary to make an estimate for a control decision. Sequential sampling, in which sampling is terminated when the pest population density is classified as either above or below a critical density, is more efficient (Onsager 1976). Sequential sampling plans for cabbage looper, *Trichoplusia ni* (Hübner), (Shepard 1973) and imported cabbageworm, *Pieris rapae* (L.), (Harcourt 1966a), have been developed for treatment decisions on fresh market cabbage. However, treatment decisions based on sequential sampling of insect pests alone can be unrealistic. If the insect pest population is within a critical range, the treatment decision is difficult, because additional variables such as present and expected growing conditions, previous crop stress, crop growth stage, expected crop value, and environmental factors that affect the pest population also must be considered. The treatment decision can be made easily on the basis of insect pest population alone if the density is either very low or very high. A sampling procedure should result in the most precise estimate when many variables must be considered in the treatment decision, but for efficiency, the estimate can be less precise when the treatment decision can be made solely on the basis of insect pest population density. An important objective in cabbage pest management is the early detection of other less common but potentially serious insect, disease, and weed pests. This requires an extensive survey of the entire field, which cannot be done if sampling is terminated quickly based on a sequential sampling decision. A need exists for an efficient sampling plan for any pest that might require control measures, but focuses on information for treatment decisions for the most important cabbage pests, *T. ni* and *P. rapae* (Andaloro et al. 1981). A sampling procedure was developed and tested based on the aforementioned objectives and the results are reported here.

Materials and Methods

Sampling occurred along a V-shaped transect, extending from one corner of the field to the middle of the opposite side and back to the other corner. One to four randomly selected cabbage plants were inspected along each of 10 equal segments of the transect. As sampling proceeded, the number of plants sampled along the next segment varied, depending on estimates of *T. ni* and *P. rapae* population density derived from previous segments. Insect pests, other than *T. ni.* and *P. rapae*, were noted on the plants examined. We conducted an extensive visual survey for diseases and weeds while walking along the transect.

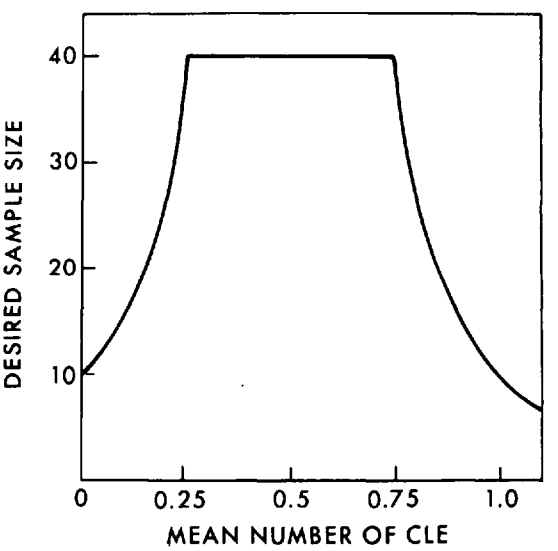

FIG. 1.—Relationship of desired sample size, number of plants to be inspected, to the mean number of CLEs per plant.

[1]Approved by the Director, N.Y. State Agric. Exp. Stn., for publication as Journal Paper No. 3402. Received for publication 15 December 1981.

[2]Dept. of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853.

[3]IPM Programs, Cornell University, Geneva, NY 14456.

0022-0493/83/0101-3905$02.00/0

Statistical Definition.

Three steps were involved in the statistical development of this sampling procedure. First, the precision desired in sampling was established as a function of the mean, or the pest population density. Second, the sample size necessary to achieve the desired precision was calculated, also as a function of the mean. The objective was to estimate the underlying mean after each segment by using previous observations and then sample at an intensity to achieve the sample size desired for the estimated mean. The third step, therefore, was to design a system to identify the number of plants that should be sampled along any segment of the transect, given the estimate of the mean from previous segments. The final estimate of pest population density is calculated as if the sample size had been fixed before sampling. This is a reasonable estimate, considering the work of Anscombe (1949).

The function relating sample precision to the mean was based on an action threshold of 0.5 cabbage looper equivalent (CLE) per plant for fresh market cabbage (Shelton et al. 1982). This technique standardizes insect damage by weighting each insect species by foliage consumption (1 CLE = 1 *T. ni* = 1.5 *P. rapae*). It was calculated that if a field is treated at this threshold throughout an average year, at least 95% of the cabbage heads would have no lepidopterous feeding damage. However, the actual treatment threshold that will result in 95% heads without feeding damage can vary from 0.25 to 0.75 CLE per plant, depending upon growing conditions and pest population pressure. Therefore, the critical region where the highest precision and the highest intensity of sampling were desired was from 0.25 to 0.75 CLE per plant. Below population densities of 0.25 CLE per plant, the precision was established by placing the upper boundary of a 95% confidence interval on the mean at the average treatment threshold of 0.5 CLE per plant; above population densities of 0.75 CLE per plant, the precision was established by placing the lower boundary of a 95% confidence interval on the mean at 0.5 CLE per plant.

The desired sample size was calculated as a function of the mean, to achieve the precision specified above. Suppose that a total of n plants are inspected and x CLEs are observed where the number of CLEs on each plant follows a negative binomial distribution with known dispersion parameter k and unknown mean m. The total number of CLEs observed has mean nm and variance $ns^2_{(m)}$, where $s^2 = m + (m^2/k)$. With a normal approximation, the fixed sample size 95% confidence interval for m is (m_1, m_2), where

$$m_1 + \sqrt{\frac{s^2_{(m_1)}}{n}} \cdot Z_{0.975} = \frac{x}{n} \qquad (1)$$

and

$$m_2 - \sqrt{\frac{s^2_{(m_2)}}{n}} \cdot Z_{0.975} = \frac{x}{n} \qquad (2)$$

If the true mean m is below 0.25, we require m_2, the upper boundary of a 95% confidence interval, to be 0.5 when $x/n = m$. Substituting into equation 2 gives the required sample size,

$$n = \frac{(Z_{0.975})^2 \, s^2_{(0.5)}}{(0.5 - m)^2} \qquad (3)$$

Similarly, if m is above 0.75 we require m_1, the lower boundary, to be 0.5 when $x/n = m$, and substituting into equation 1 gives the sample size,

$$n = \frac{(Z_{0.975})^2 \, s^2_{(0.5)}}{(m - 0.5)^2} \qquad (4)$$

Note that the full width of the desired confidence interval (m_1, m_2) varies with m, and is larger, e.g., for m = 0.75 than for m = 0.25. With k = 2, if a mean of 0.25 is used in equation 3, and a mean of 0.75 is used in equation 4, the calculated sample size in both cases is 40 plants. The region from 0.25 to 0.75 CLE per plant was where the highest precision was required, so in this region the desired sample size was fixed at 40 plants. Outside of this region, the desired sample size was established based on equation 3 or 4 (Fig. 1).

The estimate of variance used in these calculations was based on studies of the spatial distribution of the immature stages of *T. ni* and *P. rapae* by Harcourt (1961, 1965). Counts for these two pests on randomly selected cabbage plants followed the negative binomial distribution. The variance of the negative binomial is related to the parameter k, a measure of aggregation, by the formula $s^2 = m + (m^2/k)$. Harcourt proposes a common k of 1.79 for all stages of *T. ni* larvae (1966b), and 2.52 for all stages of *P. rapae* larvae (1966a). Since this procedure is based on CLEs, it was desirable to have an estimate of a common k for both species; therefore, a common k of 2, a rough approximation, was used for the present calculations.

In practice, the mean is not known when sampling begins, it must be estimated as observations are re-

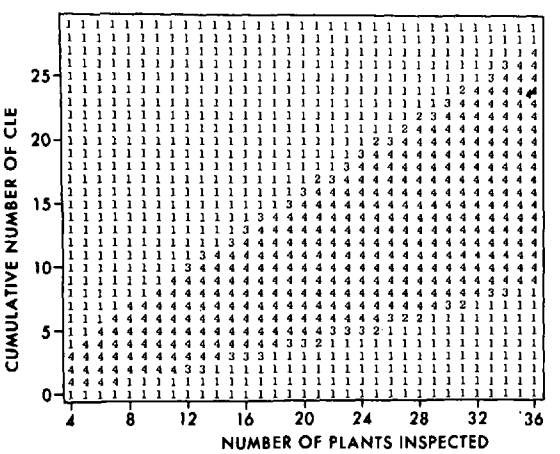

FIG. 2.—Number of plants to inspect along the next of 10 segments of a V-shaped transect to estimate pest population density by using variable-intensity sampling in a fresh market cabbage field.

corded. The transect is divided into 10 segments and a maximum of four plants, because 40 plants is the maximum sample size; and a minimum of one plant, because at least one observation is desired in each segment, is to be sampled along each segment. To give a basis for deciding the intensity along the following segments, sampling along the first segment of the transect was fixed at the maximum intensity of four plants. The number to be sampled along each successive segment after the first depends on the average number of CLEs per plant observed so far. If this number is between 0.25 and 0.75, the maximum sample size of 40 plants is indicated and four plants should be sampled; if the number is less than 0.25 or greater than 0.75, a sample size of less than 40 is indicated and the sampling intensity must be reduced at some point. To achieve approximately the required number of observations for the underlying mean, the sampling intensity along successive segments of the transect was established by finding a procedure that would give the required number of observations for any mean in the hypothetical case where each observation is exactly the mean. Table 1 gives the points, selected as guides, at which the sampling intensity changes. As an example, if each observation were 0.79, the sample size desired would be 30 plants and the first four segments of the transect would be sampled at an intensity of four plants per segment, the next three segments at three plants per segment, etc. A number of combinations of the four intensities could be used to produce a sample size of 30 plants, but the intensity combinations in Table 1 have practical merit in that the points at which the sampling intensity changes were graphed and, by interpolation, regions were established for the four sampling intensities. These regions were then used to make Fig. 2, a chart that can be used to determine the number of plants to sample along the next segment of the transect, given the number of plants already sampled and the accumulated number of CLE found, rounded to an integer. The chart does not provide the exact sample size required for all means, but it does provide the correct sample size for many means and is within three observations for the rest in the hypothetical case where each observation is exactly the mean. The distribution of the sample sizes actually achieved when observations follow the negative binomial distribution was investigated by computer simulation.

Test of Procedure

The sampling procedure was tested on observations from the negative binomial distribution by a computer simulation to: compare the estimate of the mean produced by the sample and the true mean; compare the desired sample size and the actual number of plants sampled for a given mean; and determine the frequency of a fixed sample size 95% confidence interval about the observed mean containing the true mean. A computer program was developed that generated, for a given true mean, random numbers from the negative binomial distribution with $k = 2$, and treated these numbers as counts of CLE on randomly selected cabbage plants to simulate the use of the sampling procedure field. One thousand replications of the test were used on means from 0.1 to 1.5 in increments of 0.1.

In practice, the value of k for counts of CLE is not constant, because the distribution pattern of *T. ni* and *P. rapae* varies from field to field and from day to day.Therefore, estimates produced by variable-intensity sampling were compared with estimates based on fixed sample sizes of 40 plants and 264 plants, using counts of *T. ni* and *P. rapae* taken in a commercial cabbage field near Geneva, N.Y., in July and August 1980. The field was divided into a grid with 66 equal sections. Four plants, selected at random, were inspected in each section. Counts from 10 sections along a V-shaped transect were used to estimate pest population density based on a fixed sample of 40 plants and a variable-intensity sample. This procedure was replicated over five different sampling dates in the commercial field.

Results

Computer simulations resulted in statistically acceptable performance of the procedure on the negative binomial distribution with $k = 2$ (Table 2). For means below 0.5, the average estimate was lower than the true mean; for means of 0.5 or greater, the average estimate was greater than the true mean. For means from 0.2 to 0.8, the desired sample size was greater than the average

Table 1.—Desired number of plants to be sampled along each of 10 segments of a V-shaped transect in the case where each observation of CLEs is exactly the mean

No. of segments to sample at an intensity of:				Resulting sample size (n)	Means requiring the sample size n
Four plants/ transect segment	Three plants/ transect segment	Two plants/ transect segment	One plant/ transect segment		
10	0	0	0	40	$0.25 \leq$ means ≤ 0.75
6	4	0	0	36	0.24 or 0.76
5	4	1	0	34	0.23 or 0.77
4	3	2	1	30	0.21 or 0.79
3	2	1	4	24	0.18 or 0.82
2	1	0	7	18	0.13 or 0.87
1	0	0	9	13	Means ≤ 0.06, and ≥ 0.94

Table 2.—Properties of estimates obtained from computer simulations of variable-intensity sampling

True mean	Distribution of the estimate of the true mean[a]		Desired sample size[b]	Distribution of actual sample size		Estimated probability of a fixed sample size 95% CI containing the true mean
	Mean	SD		Mean	SD	
0.1	0.08	0.01	16	17.2	5.6	0.970 (0.005)[c]
0.2	0.17	0.09	28	24.1	9.2	0.966 (0.006)
0.3	0.27	0.13	40	30.7	9.4	0.968 (0.006)
0.4	0.39	0.15	40	34.2	7.8	0.958 (0.006)
0.5	0.51	0.18	40	34.7	7.8	0.931 (0.008)
0.6	0.63	0.21	40	32.7	9.1	0.932 (0.008)
0.7	0.76	0.23	40	29.6	9.8	0.944 (0.007)
0.8	0.86	0.23	28	25.9	9.6	0.956 (0.006)
0.9	0.97	0.25	16	23.2	9.1	0.947 (0.007)
1.0	1.07	0.28	13	19.5	7.3	0.955 (0.007)
1.1	1.17	0.30	13	17.9	6.2	0.960 (0.006)
1.2	1.26	0.32	13	16.9	5.5	0.963 (0.006)
1.3	1.36	0.36	13	16.0	4.5	0.954 (0.007)
1.4	1.44	0.36	13	15.5	3.9	0.972 (0.005)
1.5	1.54	0.41	13	14.7	3.1	0.960 (0.006)

[a]Based on 1,000 replications for each mean.
[b]According to Fig. 1.
[c]SD of the estimate.

sample size resulting from the procedure; for means below 0.2 and above 0.8, the desired sample size was smaller than the average sample size resulting from the procedure. The frequency of a fixed sample size 95% confidence interval about the observed mean containing the true mean ranged from 0.931 to 0.972.

A comparison of estimates of pest population density derived from a variable-intensity sample and estimates derived from fixed sample sizes of 40 plants and 264 plants in a commercial cabbage field revealed only small differences (Table 3). The number of plants sampled by the variable-intensity procedure ranged from 13 to 40.

Discussion

The precision and the reduction of average sample size when using variable-intensity sampling make it a useful tool in pest management programs. The computer simulation demonstrated that the procedure worked well when observations follow the distribution for which it was designed. Changes in the value of k, which are to be expected in the field, made small differences in the precision achieved when using the procedure.

Variable-intensity sampling has advantages over conventional methods. It is more efficient than procedures that utilize a fixed sample size, because the effort required for a precise estimate of pest population density is expended only when necessary to make a management decision. It is more useful than a sequential sampling plan for cabbage pest management, because it requires that the procedure be implemented within a designated sampling pattern, the V-shaped transect. This requires that the field be covered by a visual survey and insures that plants are inspected throughout the field, minimizing the chances of missing an isolated pocket of infestation of any of the insect pests that could only be found by inspecting plants. Sequential sampling along a V-shaped transect could require the inspection of plants to be either terminated before or continued after the transect had been covered. Variable-intensity sampling does not require an exact action threshold, only a range of pest densities where treatment may be necessary, depending on additional variables. The result of the variable-intensity sampling procedure is not a treatment decision, as with sequential sampling; rather, it is information required to make a treatment decision. Possibilities for further improvements in the sampling procedure include a better definition of the precision needed in the estimate of pest population density, se-

Table 3.—Comparison of estimates of pest population density in a commercial cabbage field near Geneva, N.Y., in July and August 1980, based on variable-intensity sampling and fixed sample sizes of 264 and 40 plants

Date	Fixed sample size of 264 plants[a]		Fixed sample size of 40 plants[b]		Variable-intensity sample[b]		
	Mean	s^2	Mean	s^2	Mean	s^2	No. of plants sampled
18-VII-80	0.67	0.70	0.75	0.81	0.84	0.86	31
25-VII-80	0.71	0.69	0.60	0.54	0.59	0.56	34
1-VIII-80	0.05	0.28	0.03	0.16	0.00	0.00	13
15-VIII-80	2.50	1.99	2.20	1.60	2.31	1.49	13
22-VIII-80	0.22	0.50	0.25	0.54	0.25	0.54	40

[a]Samples taken from a uniform grid pattern throughout the field.
[b]Samples taken along a V-shaped transect within the field.

lection of a transect that would lower the sample variance, and an improved value for a common k for *T. ni* and *P. rapae*.

Acknowledgment

We express our gratitude to A. J. Sawyer and D. S. Robson for their helpful comments and suggestions toward the development of this procedure and to the Cornell University Vegetable IPM Programs.

REFERENCES CITED

Andaloro, J. T., A. M. Shelton, and C. J. Eckenrode. 1981. Seasonal abundance of lepidopterqus larvae in commercial cabbage fields. Environ. Entomol. 11: 144–146.

Anscombe, F. J. 1949. Large sample theory of sequential estimation. Biometrika 36: 455–458.

Harcourt, D. G. 1961. Spatial pattern of the imported cabbageworm, *Pieris rapae* (L.), on cultivated Cruciferae. Can. Entomol. 93: 945–952.

1965. Spatial pattern of the cabbage looper, *Trichoplusia ni* (Hübner), on crucifers. Ann. Entomol. Soc. Am. 58: 89–94.

1966a. Sequential sampling for the imported cabbageworm, *Pieris rapae* (L.), Can. Entomol. 98: 741–746.

1966b. Sequential sampling for use in control of the cabbage looper on cauliflower. J. Econ. Entomol. 59: 1190–1192.

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

Shelton, A. M., J. T. Andaloro, and J. Barnard. 1982. Effects of cabbage looper, imported cabbageworm, and diamondback moth on fresh market and processing cabbage. J. Econ. Entomol. 75: 742–745.

Shepard, M. 1973. A sequential sampling plan for treatment decisions on the cabbage looper on cabbage. Environ. Entomol. 2: 901–903.