

Kansas State University Libraries

**New Prairie Press**

---

Conference on Applied Statistics in Agriculture

1990 - 2nd Annual Conference Proceedings

---

## A SIMULATION STUDY ON THE RELATIONSHIP BETWEEN THE ABUNDANCE AND SPATIAL DISTRIBUTION OF INSECTS AND SELECTED SAMPLING SCHEMES

J. L. Willers

D. L. Boykin

J. M. Hardin

T. L. Wagner

R. L. Olson

*See next page for additional authors*

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

---

### Recommended Citation

Willers, J. L.; Boykin, D. L.; Hardin, J. M.; Wagner, T. L.; Olson, R. L.; and Williams, M. R. (1990). "A SIMULATION STUDY ON THE RELATIONSHIP BETWEEN THE ABUNDANCE AND SPATIAL DISTRIBUTION OF INSECTS AND SELECTED SAMPLING SCHEMES," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1430>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact [cads@k-state.edu](mailto:cads@k-state.edu).

---

**Author Information**

J. L. Willers, D. L. Boykin, J. M. Hardin, T. L. Wagner, R. L. Olson, and M. R. Williams

A SIMULATION STUDY ON THE RELATIONSHIP BETWEEN THE ABUNDANCE AND SPATIAL DISTRIBUTION OF INSECTS AND SELECTED SAMPLING SCHEMES

J. L. Willers<sup>1</sup>, D. L. Boykin<sup>2</sup>, J. M. Hardin<sup>3</sup>, T. L. Wagner<sup>1</sup>, R. L. Olson<sup>1</sup>,  
and M. R. Williams<sup>4</sup>

ABSTRACT

During the development of a Bayesian approach to estimate insect population abundance, it was necessary to compare not only the reliability of Bayesian estimates, but to also compare these estimates to those obtained by traditional methods employed by entomologists. To facilitate these comparisons it was necessary to use simulated fields apportioned into quadrats where conditions representative of insect abundance and dispersion are modeled. Thus, a simulation model was developed using SAS to derive example insect populations from which samples could be drawn. The negative binomial distribution was used to simulate the proportion of infested plants ( $p$ ) with various degrees of clustering ( $k$ ) for specified quadrat sizes. Another component varies sample parameters which represent the total number of plants sampled per field, the number of plants sampled per quadrat, and thus the number of quadrats sampled per field.

1. INTRODUCTION

A substantial proportion of the production costs in cotton is for insect control, averaging in 1988 nearly \$30/acre throughout the cotton belt (Head 1989). With continued trends toward the reduction of the number of registered insecticides and increased public pressure for environmental concerns, control costs are expected to increase. Conversely, when damaging densities of insects are not controlled, losses in net profit and fiber quality occur. Thus, it becomes increasingly important that adequate methods for estimating the abundance of insect pests be developed so that insecticides are applied only when necessary. Additionally, the continuing development of expert systems (Jones 1989; Olson et al. 1990), means that not only these programs, but the human experts who build these systems, need dependable information to provide reliable recommendations. For example, an input prompt from the rule-based expert-system, CIC-EM (Bowden et al. 1990), reads: "Select the parameter that best indicates the number of *Heliothis* larvae per acre found in terminals (Low: 0 to 1500, Medium: 1500 to 4000, High: Above 4000)". At typical planting densities of 40,000 plants per acre, the selection of "High" corresponds to an infestation rate of  $\geq 10$  per cent.

Sampling insect populations is a traditional topic among entomologists (Ruesink 1980; Sterling and Pieters 1979; Wilson et al.

---

<sup>1</sup>USDA, ARS, Crop Simulation Research Unit, Mississippi State, Mississippi

<sup>2</sup>USDA, ARS, Delta State Research Center, Stoneville, Mississippi

<sup>3</sup>University of Alabama-Birmingham, Birmingham, Alabama

<sup>4</sup>Mississippi Cooperative Extension Service, Mississippi State, Mississippi

1989). Despite the abundance of published sampling plans, demands for improved methods of sampling are still found in recent publications (Frisbie et al. 1989; Reissig et al. 1989). Therefore, efforts were directed toward exploring the possibility of using Bayesian concepts to develop robust, efficient sampling plans.

In the biological sciences, including entomology, the Bayesian technique has recently been addressed (Gazey and Staley 1986; Nyrop et al. 1986; Plant and Wilson 1985). Some attractive features of Bayesian methods, germane to this study, include: (1) the immediate and automatic use of all available data, and (2) using the decision maker's "degree of belief" in how an unknown parameter changes with increasing information (Box and Tiao 1973; Johnson 1977). Further, according to Gazey and Staley (1986), all quantities are of two kinds- those known and those unknown to the person making the inference. If one is able to adopt the position that the uncertainty of the unknown quantity can be described by a probability distribution, a formal procedure for solving inference problems is obtained. Building upon these concepts, a useful approach is provided for the evaluation and update of sparse information. A situation typically encountered in the estimation of insect populations in cotton is used for illustration.

#### 1A. Illustrative Problem:

Consider that for the purpose of commercial monitoring (Wilson et al. 1989), an entomologist determines that an examination of 100 plants in a 50 acre cotton field is sufficient for making a judgement on the abundance of *Heliothis* eggs. In his or her opinion, a sample of this size represents a manageable sample (Schmitt 1969), especially in comparison to the 2,000,000 or more plants that are present within the field. Clearly, if all 100 plants have or do not have eggs, confidence is high about the condition of the remaining hundreds of thousands of plants left unexamined. But, if 5 to 35 of 100 plants are found to have eggs, confidence about the character of the rest of the field is less certain. Thus, these points, in addition to the decision rule used to determine if an insecticide should be used, represent some of the information known prior to sampling.

Before sampling, there is some information which is unknown. The principle unknown quantity,  $R$ , is the probable number of plants that can be found infested with *Heliothis* eggs. Secondly, although 100 plants is the declared finite sample size ( $N$ ), it may not be possible to always achieve a sample of this size, thus the actual sample size for a field can vary. Also, the number of distinct locations at which plants are examined can vary from field to field.

In many extant sampling plans, the sample size relative to the population size of plants at risk to insect attack is extremely sparse. Further, most sampling schemes do not utilize information about the number of locations (sub-samples) in a field examined for insects to obtain an estimate of the field infestation rate ( $p$ ). Within these constraints, an iterative Bayesian approach, which uses the known and unknown information, can be used to infer the probable proportion of plants at risk to insect attack. This inference constitutes the basis for the management decisions on the field.

The current solution approach is developed from concepts presented in Berger (1985), Gazey and Staley (1986) and Schmitt (1969). The selected probability model is the hypergeometric distribution, a discrete

distribution applicable to binomial (presence or absence) sampling without replacement. The form used here, corresponding to that used by Schmitt (1969), is:

$$P(R|N, n, r) = \frac{\binom{R}{r} \binom{N-R}{n-r}}{\binom{N+1}{n+1}} \quad (1)$$

Note that the denominator must be modified because the domain for  $R$  and  $r$  include the alternative, 0. For example,  $r \leq 0, 1, 2, \dots, n$ . The domain for  $R$ , and the definition of notation is presented below.

The data utilized for this illustrative example are from an actual scouting report. The known information used to parameterize eq. (1) are the following quantities: (1) the declared, finite sample size ( $N = 100$ ), (2) the sub-sample sizes ( $n_i = 9, 14, 17$  or 17 plants) from four locations ( $i$ ), and (3) the corresponding number of events of interest,  $r$ , (here, the presence of at least one *Heliothis* egg upon a plant) observed in each respective sub-sample ( $r_i = 0, 1, 1, \text{ or } 3$ ). The unknown information, represented by the parameter  $R$ , is a vector of 101 competing, but equally likely, alternatives a priori ( $R = 0, 1, 2, 3, \dots, 100$ ). Currently, this initial vector is a non-informative, uniform prior (Gazey and Staley 1986; Schmitt 1969), which is the reciprocal of the finite number of alternatives. The task is to calculate the degree of belief, portrayed by the posterior distribution [ $\text{Prob}(R|N, n, r)$ ], for each of these alternatives given the information obtained from the samples (Box and Tiao 1973; Schmitt 1969).

The rule of Bayes (Box and Tiao 1973) can be given by a recursive formula (Gazey and Staley 1986):

$$\text{Posterior distribution}_{(i)} \approx \text{Likelihood}_{(i)} \times \text{Prior distribution}_{(i-1)}. \quad (2)$$

Since there are four locations ( $i = 1, 2, 3, 4$ ) from which samples were acquired, there will be four iterations of the sequential Bayesian algorithm. To properly iterate, the information from the locations must be rank ordered by  $r$  within  $n$ , and the  $\sum n_i \leq N$ . [This algorithm is not to be confused with the Bayesian sequential sampling method of Plant and Wilson (1985)]. The non-informative prior,  $\text{Prior}_{(0)}$ , is modified by the likelihood calculated from the first sample ( $i = 1$ ;  $\text{Prob}(R|100, 9, 0)$ ). Continuing the algorithm, the first posterior distribution becomes the prior to be updated by the likelihood of the second sample, and so on until the final posterior distribution is calculated ( $i = 4$ ). The sequences of distributions, including the non-informative prior, are graphed in Fig. 6. The final sequence is shaded.

The final posterior distribution represents the entomologist's knowledge about the status of the field. The mode of this distribution is 9, which is approximately equal to the ratio of the sums,  $\sum r_i / \sum n_i$ , or 5/57. From the calculated probabilities, it is possible to construct the cumulative distribution function, and define the standardized probability region (Gazey and Staley 1986; Johnson 1977), the limits  $R_1$  and  $R_2$ , which define the interval:

$$\text{Prob}(R_1 < R < R_2) = 0.95. \quad (3)$$

Of primary interest is the information provided by the pattern of the posterior distribution sequences (Gazey and Staley 1986). For this illustration, it is observed that three of the four posterior distributions have modes  $\leq 5$ . The mode of the final posterior is  $> 5$ , because the likelihood of the fourth and final sample (not shown on the graph) is centered about  $R = 18$  (or  $3/17$ ), shifting the final posterior to centralize about  $R = 9$ . Suppose for the purpose of illustration, that the decision rule (or threshold) to use an ovicide was 10 per cent ( $R \geq 10$ ). A cautious decision maker might choose to spray, since the mode and the standardized probability interval of the final posterior are close to and include values of  $R \geq 10$ . However, the pattern of the four sequences suggest he should choose not to spray.

In order to understand how the pattern of the iterative sequences supply the evidence to choose the "No spray" decision, additional information needs to be elaborated. The eggs of *Heliothis*, as well as young larvae, are typically dispersed as clusters (or aggregates) in the field (Allen et al. 1972; Pieters and Sterling 1973; Wilson and Room 1983). Thus, the three likelihoods that center over low values of  $R$ , and the one likelihood that centered over  $R = 18$ , reflect a clustered, but sparse distribution for the insect population currently present in this field. In fact, for this field, 12 additional locations, representing 204 more plants were examined, and only one more egg was found. The additional (over)sampling, in this instance, confirmed the conclusion provided by the spatial patterns of the sequences which used a much smaller sized sample. By recursively utilizing sub-sample data with Bayes' theorem, additional information that would otherwise be discarded is available to the decision maker. These results provide the motivation for the simulation experiment described in this paper.

## 2. SIMULATION OF INSECT DENSITY AND SPATIAL ABUNDANCE

The purpose of the simulation experiment is to explore the following hypothesis suggested by the previous example: Is it possible to use the distribution of the likelihoods (or posterior distributions) of an iterative Bayesian algorithm to infer if the relative estimate of population abundance ( $p$ ) is robust, and if the insect population is clustered or randomly dispersed? Another application of the simulation analysis is to provide direction for developing a sampling program to be used in the field.

The topic of spatial pattern analysis in ecology has been addressed by many authors, including Ludwig and Reynolds (1988), Pielou (1977) and Southwood (1978). In many of these analyses, the spatial patterns are described by probability distributions. In this study, the negative binomial distribution, as described by Anscombe (1949), is used to model the spatial pattern of *Heliothis* eggs among quadrats. The negative binomial distribution describes the probability ( $P_{R'}$ ) of having  $R'$  infested plants in a quadrat containing  $N'$  plants:

$$P_{R'} = \binom{k+R'-1}{R'-1} \left[ \frac{m}{m+k} \right]^{R'} \left[ \frac{k}{m+k} \right]^k \quad (4)$$

where

$k$  = clustering parameter,

$m = N' * p$ , the average number of infested plants per quadrat, and  
 $p$  = average proportion of infested plants per field (grid).

Also,

$$\begin{aligned} E(R') &= m, \\ \text{Var}(R') &= m + m^2/k, \\ k &= m^2/(\text{Var} - m). \end{aligned}$$

Quadrat size is determined by the number of plants ( $N'$ ) contained in a cell, based upon a density of 3 plants per foot of row, with 40 in. row spacings. The quadrat determines the area in which clustering is expected to occur; however, the point pattern of infested plants within quadrats is assumed to be randomly distributed.

Using SAS,  $P_{R'}$  may be obtained by  $P_{R'} = \text{PROBNEGB}(PP, k, R')$ , where  $PP = k/(m+k)$ . The SAS programs used to perform these simulations are available upon request. The computer simulation has two parts: (1) To simulate solid planted (that is, there are no skip rows) cotton fields and randomly place insects in the field based on given parameters of the negative binomial (Figs. 1-5) and (2) To randomly select a given number of plants from the simulated field, and interpret the information taken from the sampled plants by several methods (Fig. 8 and Table 1). The insect densities selected are 1, 5, 10, 15, 20 and 25 per cent, with various levels of clustering [high ( $k = 1$ ), medium ( $k = 2, 3$ ), and low ( $k = 50$ )] for different quadrat sizes of 50, 100, 150, or 200 plants. Additional parameters determining sampling characteristics, and which attempt to approximate the behavior of a cotton scout are:  $NS$  = Total plants checked,  $S$  = No. of stops (quadrats) per field, and  $NP$  = No. of plants checked at each stop.

Other constraints currently placed on the simulation model are that only one spatial distribution represents an entire field, each stop occurs in a different quadrat and there is no bias in the selection of the quadrat or of the plants within it. Further, the information obtained from the sampled plant is limited to either a "yes or no" rating, and there is no measurement (scouting) error since an infested plant, if sampled, is always rated as "yes". These constraints describe what has been done up to this point, although there are other constraints that could be considered to improve the realism of the model. For example, cotton fields are usually quite large, and the decision to spray or not spray will usually apply to the whole field. In reality there are probably several spatial distributions and infestation rates within a field, which complicates the making of management decisions. Later studies will address this and other issues.

### 3. SIMULATION AND SAMPLING RESULTS

A simulated tract of a cotton field is shown in Figure 1. The dimensions of this 7 x 18 grid are about 50 by 150 feet, with the rows running horizontally. Each rectangle represents a quadrat of  $N' = 50$  plants, or an area of 2 rows x 8.333 feet. This plot represents the dispersion of infested plants when the negative binomial is parameterized with a  $k$  value of 1 and an overall infestation rate ( $p$ ) of 10 per cent. The symbol within each quadrat depicts the number of infested plants where  $R'$  ranges from 1 to 50 plants. While Fig. 1 depicts a field with a high

level of clustering ( $k = 1$ ), Fig. 2 is similar, but with less clustering of infested plants ( $k = 50$ ) within the grid. The change from a quadrat size of 50 to 200 plants, for a grid equivalent in area to that of Fig. 1, is illustrated by Fig. 3 for  $k = 1$  and  $p = 0.10$ . Note the adjustments in the symbol scale.

For either field in Figs. 1 and 2, a 10 per cent infestation rate is an expectation of 5 infested plants per quadrat. However, the clustering parameter,  $k$ , determines different dispersion patterns among the grids. When there is not much clustering (Fig. 2), most quadrats tend to have about 5 infested plants, with a few not having any infested plants. When there is clustering, most quadrats are without infested plants, and the frequency of quadrats with 11 or more infested plants is greater. A quadrat with 11 or more infested plants is arbitrarily called a "hotspot". These are quadrats where the number of infested plants is more than twice the average for the field ( $R' \geq 2*m$ ). Analogous sites occur within a field in practice, and often are the same places year after year.

The probability plots of Figs. 4 and 5 show distributions for either the negative binomial with different quadrat sizes and levels of  $k$  (Fig. 4), or with different levels of  $k$  and  $p$  for a constant quadrat size of 50 plants (Fig. 5). These plots permit the evaluation of how  $k$  relates to changes in quadrat size ( $N'$ ) and infestation rate ( $p$ ). For example, in Fig. 4A,  $R'$  ranged from 0 to as many as 25 out of 50. When  $k = 50$ , the bell shaped distribution shown approximates the binomial distribution, where on the average ( $m$ ) 5 infested plants will occur per quadrat. When the value of  $k$  is decreased, the variance of  $R'$  increases and the distributions become skewed, with more frequent occurrences of quadrats with fewer or greater numbers of infested plants than expected on average. (However, for a 1 per cent infestation rate, these and the following generalizations do not apply.)

With a 10 per cent infestation rate, it is observed that the frequency of quadrats with none or few infested plants ( $R' \leq 0.2*m$ ; that is, the area of the left-hand tail) is quite variable when quadrat size is increased (Fig. 4). With a constant infestation rate of 10 per cent, the distributions are pulled to the right when the quadrat size is increased from 50 to 200 plants (Figs. 4A-D). This shift decreases the probability that a quadrat with none to few infested plants will occur. For the smallest value of  $k$  ( $= 1$ ), the rate of decrease in the frequency of sparsely infested quadrats is much less than for slightly larger values of  $k$  ( $= 2, 3$ ). Comparable changes in the frequency of occurrence of sparsely infested quadrats result when both  $k$  and  $p$  are varied, but when quadrat size is held constant (Fig. 5). For example, by using the tabulated frequencies (not shown) which correspond to the distributions shown in Fig. 5, the percentage of quadrats with  $R' \leq 0.2*m$  is calculated to range from 38.8 to 24.9 per cent, when  $k = 1$  and  $p$  varies from 5 to 20 per cent. But, for either situation with a much larger clustering parameter ( $k = 50$ ), the frequency of quadrats with  $R' \leq 0.2*m$  rapidly approaches zero (Figs. 4 and 5).

A different pattern of behavior is observed in Figs. 4 and 5 for the area which corresponds to the right-hand tail. As opposed to the highly variable left-hand tail, the right-hand tail is quite constant. The frequency of occurrence of hot spots is somewhat stable when the dispersion pattern among quadrats is clustered despite changes in quadrat size or infestation rate. For example, for each quadrat size and when  $k = 1$ , approximately 13 per cent of the quadrats are hot spots (Fig. 4A-D).



Similarly, the frequency of quadrats categorized as hot spots is fairly uniform at roughly 13 per cent, when the infestation rate is varied from 5 to 20 per cent and quadrat size is fixed ( $N' = 50$ ; Figs. 5A-D). The right-hand tail is most variable only when the clustering parameter is larger ( $k = 50$ ). When this is the case, the probability of observing a quadrat where  $R' > 2*m$ , is not constant and ranges from 1.8 to zero per cent (Fig. 4), or from 4.6 to 0.4 per cent (Fig. 5). The point where  $R' \geq 2*m$  is indicated by a vertical reference line.

The cumulative plots of  $P_R$ , provide the evidence which answer the hypothesis concerning the use of likelihoods or posterior distributions to best estimate insect abundance. The cumulative distribution functions (CDF) for a moderately small clustering parameter ( $k = 2$ ), small quadrat size ( $N' = 50$ ) and when  $p$  varies from 1 to 25 per cent illustrates how (Fig. 7). These plots indicate, by subtraction, the probability ( $F_{R'}$ ) of observing a particular  $R'$  for each value of  $p$ . For example, when  $p = 0.05$ , the probability  $R' = 20$  is zero, since  $F_{20} - F_{19} = 1.0 - 1.0 = 0.0$ . Generally, higher values of  $R'$  become increasingly probable as  $p$  is increased. The practical importance of these results are demonstrated by the following example.

Using simulation, samples are obtained from two fields which represent infestation rates of 5 and 20 per cent, from which 20 plants are examined from each of 5 quadrats ( $N = 100$ ,  $N' = 50$ ,  $k = 2$ ). Using a FORTRAN program, the likelihoods ( $\text{Pr } R$ ) for each sub-sample are calculated and plotted (Fig. 8). The final posterior is represented by a single vertical line with probability 1, since the  $\sum n_i = N$ . The corresponding cumulative distribution function of the parametric infestation rate is displayed above each plot of the likelihoods. As displayed, there is considerable difference between the positions of the likelihoods between the two fields. For a 5 per cent infestation, the observance of a likelihood which would center over  $R = 20$  is an extreme possibility, since the cumulative plot is beginning to plateau at this point. Whereas, at 20 per cent, the observance of a likelihood which centers over  $R = 20$ , or even greater, is quite probable. Thus, it is confirmed that low infestation rates yield likelihoods which aggregate among low values of  $R$ , as was observed for the posterior distributions of Fig. 6. The likelihoods obtained from higher infestation rates are broadly distributed among higher values of  $R$ .

These plots (Figs. 4, 5, 7 and 8) also imply how the sampling of one or more hot spots could result in the decision to use an insecticide when it is not needed. On the other hand, one could sample by chance one or more quadrats where insects do not occur and conclude that no control is needed, but the proper decision is to initiate control. The Bayesian algorithm does improve decision making, but is still susceptible to failing to initiate control when necessary. This situation is shown by the following example.

In Table 1, twenty-four simulations, in which 20 plants were examined from each of 5 quadrats, provide the data for comparison. For simple random sampling the decision to not spray was chosen whenever the estimated infestation rate ( $p'$ ) from the five quadrats was  $\leq 8$  per cent (6 occurrences). The decision to spray was chosen 14 times when the estimate was  $\geq 12$  per cent. For intermediate estimates,  $8 \leq p' \leq 12$ , the decision was classed as undecided 4 times. For the Bayesian sampling plans, rules applied to the sub-sample estimates ( $m'$ ) and the final posterior ( $p'$ )

Table 1. Comparisons of two Bayesian estimators to simple random sampling when  $p = 0.15$ ,  $N' = 50$ ,  $k = 1$ , and the decision level is 0.10.

Decision	Method of Analysis		
	Bayes 1	Bayes 2	Random
No Spray	9	6	6
Spray	15	18	14
Undecided	--	--	4

comprised the decision criteria. The Bayesian approaches did not yield any cases where the decision was undecided. As an example of what these rules are, they read for Bayes 1: "If ( $p' \geq 0.12$ ) or ( $0.08 \leq p' \leq 0.12$ ) and the number of quadrats with ( $m' \geq 0.12$ )  $> 3$ ), then Spray". The use of rules slightly improved the number of correct decisions from 14 to 15 when compared to random sampling. These rules correspond to the behaviors of the distributions as observed to occur in Figs. 6 and 8.

Adding more sophistication to these rules, which we call the Bayes 2 approach, provides an improvement of 3 to 4 more correct decisions over the other two approaches. Notice however, between the Bayes 2 and the simple random sampling plan, that the number of no spray decisions is the same. These decisions are equal in number because neither sampling plan can guard against drawing by chance a sequence of quadrats where  $m \ll 0.15$ . This situation results in management error, and is another area for further study.

#### 4. CONCLUSIONS

Future plans will consider a Bayesian estimator which uses an informed, rather than the uninformed, uniform prior used here. This prior would describe the population distribution of the minimal density for which the use of an insecticide, in the judgment of the entomologist or farm manager, is justified. Accordingly, the sample information from the field will be evaluated in comparison to this informative prior. Based on the results of our simulations, the Beta distribution would be an appropriate distribution to consider for use as an informative prior. The choice of the Beta distribution is mathematically convenient, since it is a natural conjugate to the negative binomial distribution.

The making of decisions with sparse data is difficult, especially when, by chance, quadrats which are sparsely or highly infested are sampled in sequence and bias the estimate of mean abundance. Further investigations using simulation, and practical experience obtained from actual use of the Bayesian algorithm, are expected to yield progress in minimizing the effects of these biases. Notwithstanding, within a Bayesian context it is reasonable to use the sub-sample information from randomly chosen quadrats to estimate the dispersion traits and abundance of insect populations in cotton.

#### ACKNOWLEDGEMENTS

We thank Bruce Mackey for his assistance in developing the use of the negative binomial distribution to model dispersion patterns. Appreciation

New Prairie Press

<https://newprairiepress.org/agstatconference/1990/proceedings/5>

is also expressed to Sam Turner for his assistance in the development of the FORTRAN program for the hypergeometric distribution. This paper is MAFES publication no. PS-7568.

#### LITERATURE CITED

- Allen, J., D. Gonzalez, and D.V. Gokhale. 1972. Sequential sampling plans for the bollworm, *Heliothis zea*. *Environ. Entomol.* 1: 771-779.
- Anscombe, F. J. 1949. The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics* 5: 165-173.
- Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer-Verlag, New York.
- Bowden, R.O., R.G. Luttrell, and L.G. Brown. 1990. Cotton Insect Consultant for Expert Management (CIC-EM): An expert system for managing cotton insects. Working Paper No. 90-01, Department of Industrial Engineering, Mississippi State, Mississippi.
- Box, G. E. P. and G. C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Addison-Wesley, Reading, Massachusetts.
- Frisbie, R.E., K.M. El-Zik, and L.T. Wilson. 1989. The future of cotton IPM. *In Integrated Pest Management Systems and Cotton Production*, (R. E. Frisbie, K. M. El-Zik and L. T. Wilson, eds.). Wiley and Sons, New York. pp. 413-428.
- Gazey, W. J. and M. J. Staley. 1986. Population estimation from mark-recapture experiments using a sequential Bayes algorithm. *Ecol.* 67:941-951.
- Head, R. B. 1989. Cotton insect losses - 1989. *Proc. Beltwide Cotton Production Res. Conf.*, Nashville, TN. pp. 193-197.
- Johnson, D. H. 1977. Some Bayesian statistical techniques useful in estimating frequency and density. U. S. Dept. of Int., Fish and Wildlife Serv., Special Scientific Report-Wildlife No. 203. Washington, D. C.
- Jones, P. 1989. Agricultural applications of expert systems concepts. *Agric. Systems* 31: 3-18.
- Ludwig, J.A. and J.F. Reynolds. 1988. *Statistical Ecology, A Primer on Methods and Computing*. John Wiley and Sons, New York.
- Nyrop, J.P., R.E. Foster, and D.W. Onstad. 1986. Value of sample information in pest control decision making. *J. Econ. Entomol.* 79: 1421-1429.
- Olson, R. L., T. L. Wagner, and J. L. Willers. 1990. Framework for modeling uncertain reasoning in ecosystem management. I. Background and theoretical considerations. *AI Appl. Nat. Res. Man.* (in press).
- Pielou, E.C. 1977. *Mathematical Ecology*. John Wiley and Sons, New York.  
New Prairie Press  
<https://newprairiepress.org/agstatconference/1990/proceedings/5>

- Pieters, E.P. and W.L. Sterling. 1973. Inferences on the dispersion of cotton arthropods in Texas. *Environ. Entomol.* 2: 863-867.
- Plant, R. E. and L. T. Wilson. 1985. A Bayesian method for sequential sampling and forecasting in agricultural pest management. *Biometrics* 41: 203-214.
- Reissig, H., R. C. Seem, and J. Nyrop. 1989. Integration of pesticide use into IPM systems. *Proc. National Integrated Pest Management Symposium/Workshop. Las Vegas, Nevada.* pp. 96-110.
- Ruesink, W.G. 1980. Introduction to sampling theory. *In Sampling Methods in Soybean Entomology*, (M. Kogan and D.C. Herzog, eds). Springer-Verlag, New York.
- Schmitt, S. A. 1969. Measuring uncertainty: An elementary introduction to Bayesian statistics. Addison-Wesley, Reading, Massachusetts.
- Southwood, T. R. E. 1978. Ecological methods with particular reference to the study of insect populations. Halsted Press, John Wiley & Sons, New York.
- Sterling, W.L. and E.P. Pieters. 1979. Sequential decision sampling. *In Economic Thresholds and Sampling of Heliothis Species on Cotton, Corn, Soybeans and Other Host Plants. Southern Coop. Series Bull. No. 231. Oklahoma Agric. Expt. Sta., Oklahoma State Univ., Stillwater.* pp. 85-101.
- Wilson, L.T. and P.M. Room. 1983. Clumping patterns of fruit and arthropods in cotton, with implications for binomial sampling. *Environ. Entomol.* 12: 50-54.
- Wilson, L. T., W. L. Sterling, D. R. Rummel and J. E. DeVay. 1989. Quantitative sampling principles in cotton IPM. *In Integrated Pest Management Systems and Cotton Production*, (R. E. Frisbie, K. M. El-Zik and L. T. Wilson, eds.). Wiley and Sons, New York. pp. 85-119.

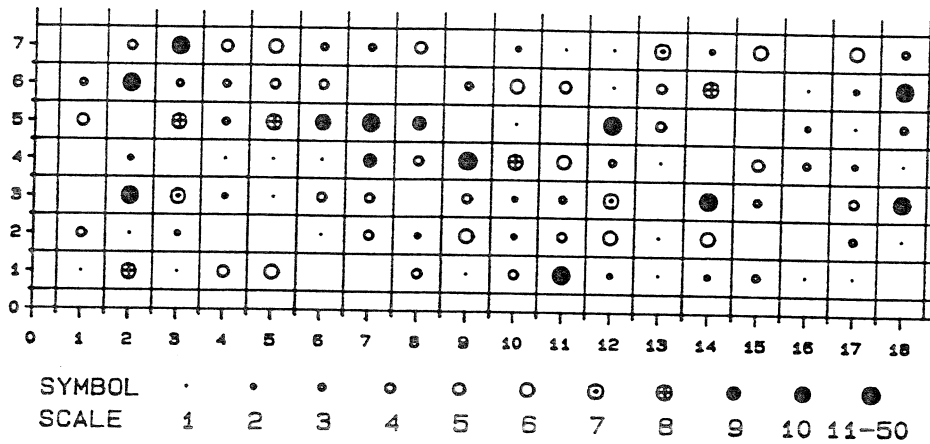


FIG. 1. Infested plants/quadrat, where  $p=0.1$ ,  $k=1$ ,  $N'=50$ .

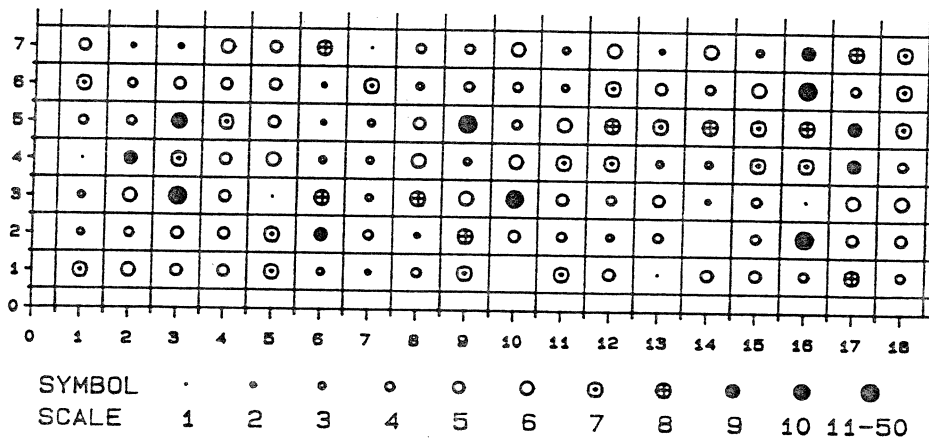


FIG. 2. Infested plants/quadrat, where  $p=0.1$ ,  $k=50$ ,  $N'=50$ .

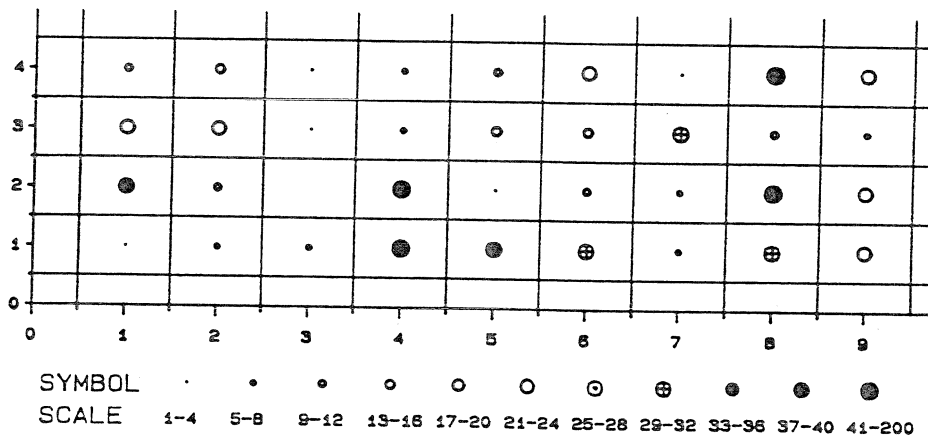


FIG. 3. Infested plants/quadrat, where  $p=0.1$ ,  $k=1$ ,  $N'=200$ .

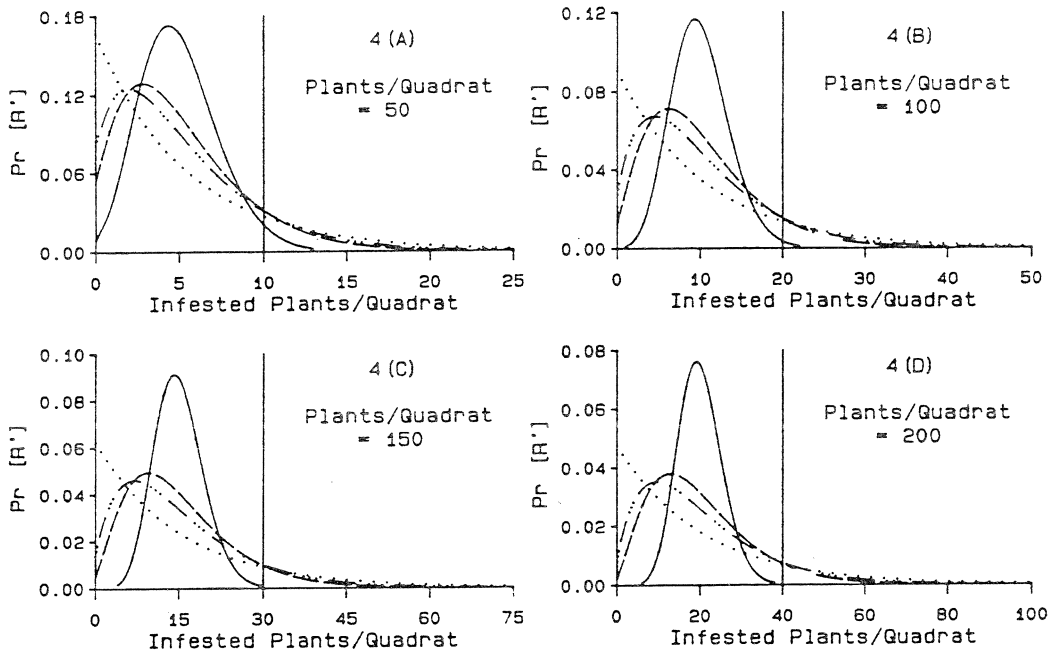


Fig. 4. Probability plots for  $p=0.1$ , variable quadrat size, and for several values of  $k$ : ..... 1, - · - · - 2, - - - - 3, ——— 50 .

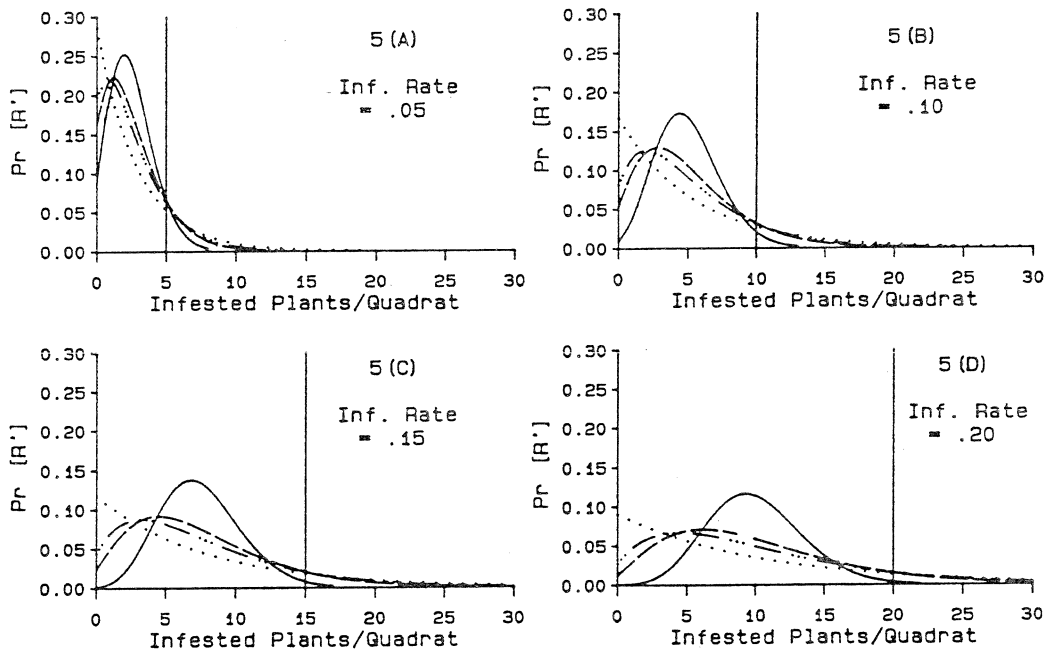


Fig. 5. Probability plots for a quadrat size of 50 plants, variable infestation rates and for several values of  $k$ : ..... 1, - · - · - 2, - - - - 3, ——— 50 .

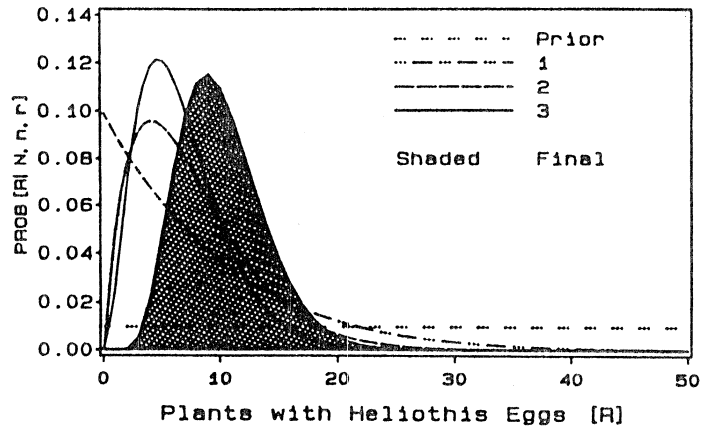


Fig. 6. Sequences of posterior distributions.

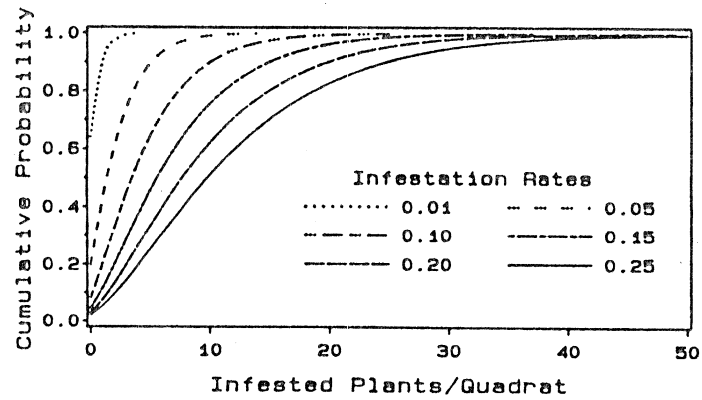


Fig. 7. Cumulative probabilities of the occurrence of quadrats with different numbers of infested plants for  $k=2, N'=50$  and selected values of  $p$ .

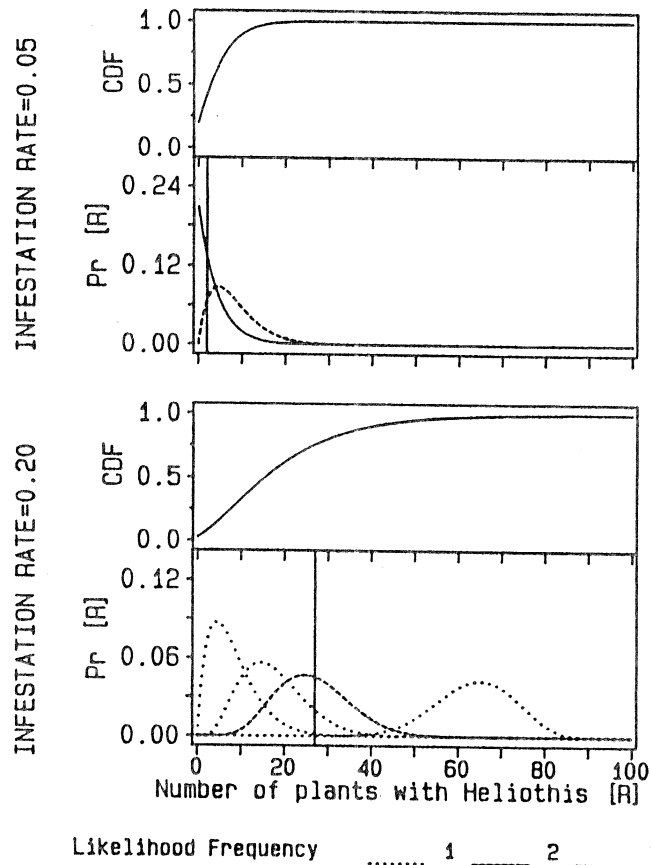


Fig. 8. Plots which compare the sample likelihoods to the corresponding cumulative distribution at two infestation rates, with  $k=2, N'=50$ . The single vertical line is the final posterior.