The potential of insect self-marking for the interpretation of trap catch

E.P. Wileyto*

Abstract

Self-marking is an approach used to make absolute estimates of insect populations using trap data. It is relatively labour free because traps and marking stations (traps modified to mark and release insects) are placed in the population simultaneously, and the user walks away. Dynamic models describing insect movements into marking stations and traps are used to determine the numbers of marked and unmarked insects expected to be found in the trap upon return. These models have been expanded to allow the estimation of parameters related to population growth, in addition to population size. The statistical methods used to infer population size, and other population parameters, are straightforward and used extensively in ecological research. This paper reviews the self-marking models and their application.

Introduction

Knowing the size of an insect population can be very important. For someone who manages a grain storage facility, this quantity may represent the information needed to make a decision about treatment. For someone else, determining population size may be an integral part of ongoing research. Regardless of need, populations are seldom estimated, except in a relative way, because of two problems. First, trapping alone gives only relative numbers or indices, not absolute estimates of size or density. Information about 'capture probability' or 'trapping efficiency' is needed to make the leap from a relative to an absolute measure, but it is not generally available. The only hope has been to relate trap efficiency statistically to the environment where the trap is placed, so that knowing enough about weather, and the items competing with traps for an insect's attention would allow us to infer capture probability (Andow 1984; Shore and McLean 1988; Hagstrum et al. 1988; Hagstrum et al. 1988). The approach is too complex; it incorporates many errors, first in describing the relationship between environment and capture probability, and second in assessing the state of the insect's current environment. The approach also demands that a manager track many aspects of the environment in order to feed the model and get the answer. Work in the field is best kept simple.

Traditional methods of population estimation, such as mark–recapture and removal, can be used to estimate not only population size, but also rates of recruitment and mortality (see Seber 1982; White et al. 1982; Nichols 1992). The problem is that these methods are too labour intensive to be used routinely for insect monitoring, and often do not perform well when used for insect research. For mark–recapture methods, the initial marking phase may require that many thousands of insects be captured, marked, and returned to the population in good condition. Even then, getting few or no recaptures is a common occurrence. Removal methods would seem to be made for use with trapping, since they work by repeating the trapping efforts, and modelling the degree to which prior trapping depresses later trapping (White et al. 1982). Yet, the return can be small for the labour, requiring two or three sequential trap observations just to estimate population size.

The answer comes in part from using the techniques of statistical inference that are behind the traditional methods of population estimation. However, we bring them to bear against a unique class of marking and trapping models built on the assumption that the insects can do the work for us (Wileyto et al. 1994). For trapping, the observer may use any of the common trap designs on the market today, or design one for special use. Marking is accomplished by marking stations, otherwise identical to traps in structure and bait, but modified to mark and release insects back to the population. Traps and marking stations are placed in the population simultaneously, and when the traps are retrieved, they contain samples of marked and unmarked insects. The difficult work of interpreting results is kept in the laboratory, on the computer. Technology in the field is kept comparatively crude and easy to use.

Self-marking methods, like the traditional methods of population estimation, can be used to obtain estimates, not only for population size, but also for other parameters such as adult recruitment rates. The approaches are also alike in that they each consist of two parts: 1) a deterministic portion describing what is expected in the trap; and 2) a probabilistic portion, there because real trap results are, in part, a result of chance. This paper will emphasise the deterministic side. I will show how the self-marking scheme is modelled to produce expected results for a trap, and discuss the variety of additions that can be made to the model to invade further the lives of our favourite insects. I will also discuss how well they perform. Finally, I will mention some specific applications of recruitment estimates that have been articulated recently.

Estimation Models

Most methods of population estimation assume that the observer cannot find and count everyone (Seber 1982). Rather, a portion of the population is caught, leaving an unknown number at large. However, by having those caught split up into different categories, according to a specific probability model which is related to population size, an estimate can be made (Seber 1982). For instance, the mark–recapture method captures and marks a known number of animals before returning them to the population. That number, divided by the total population, is the capture probability for marked individuals during the next capture round. Likewise, the removal method assumes a constant (per capita) capture probability for each of several sequential trapping episodes, while

* USDA-ARS Stored Product Insects Research and Development laboratory, 3401 Edwin Street, Savannah, GA 31405 USA.

the population is diminished by a known number over each episode.

For either of the traditional methods, capture probabilities represent what is expected to be seen on average. The actual trapping results occur in part by chance. Closed population models of mark and recapture assume a hypergeometric sampling distribution, while removal methods assume a multinomial distribution. From there it is easy to use maximum likelihood techniques to obtain simultaneous estimates of population size and capture probability. (For information on these methods, see Seber 1982; Nichols 1992).

The self-marking approach, likewise, captures a portion of the population (Wileyto et al. 1994). The relationships between capture probabilities for marked and unmarked individuals are determined from the assumption that traps and marking stations are equally attractive, and operate at the same rate. Figure 1 is a schematic representing the flow of insects from the unmarked population (F) into the trap (C, R), under the original sampling scheme. (Category M indicates marked yet free insects.) Marking stations are used at a 1:1 ratio with traps, so that both per capita marking and trapping rates are equal to λ. The trajectories for all categories are obtained by solving for a set of differential equations:

$$\begin{aligned} E(F)' &= -2\lambda F \\ E(C)' &= \lambda F \\ E(M)' &= \lambda F - \lambda M \\ E(R)' &= \lambda M \end{aligned} \tag{1}$$

Then, after substituting p for $e^{-\lambda t}$, we can specify the probabilities of being in category C, R, or of not being in the trap:

unobserved	captured, unmarked	captured, marked
p,	$1/2(1-p^2)$,	$1/2(1-p)^2$

This set of multinomial probabilities allows the use of maximum likelihood techniques to find a population estimate $(\hat{N})$ simultaneously with the parameter p, given by:

$$\hat{N} = \frac{(C+R)^2}{2R} \tag{3}$$

which has a positive bias when the number of recaptures is small. (Bias indicates that on average, the estimate misses the true population size. Bias is virtually corrected by adopting the following version with slight modification in the denomnator (see Wileyto et al.1994):

$$\hat{N} = \frac{(C+R)^2}{2(R+1)} \tag{4}$$

The approach can be improved by using a double marking scheme, which yields increased degrees of freedom. For example, suppose that for every trap, we use two red marking stations and one blue marking station. Insects can show up in the trap as unmarked (C), or red (R), or blue (Q), or red and blue (category S is doubly marked from visiting both types of marking station). Figure 2 shows a generalisation of this scheme, which uses α red and β blue marking stations. (Using unspecified ratios allows complete flexibility in sampling scheme specification after the experiment is done.) Again, solving simultaneously a set of differential equations provides the trajectory for each of the categories in the sampling scheme, and specifies multinomial probabilities to be used in estimation. Estimation then relies on numerical methods (Pollard 1977). Further, the added degrees of freedom may be used to test the fit of the model, using a chi-squared or other goodness-of-fit test.

The closed population models mentioned above have been tested thoroughly by simulating the sampling distributions of the population estimate over a wide range of parameter values,

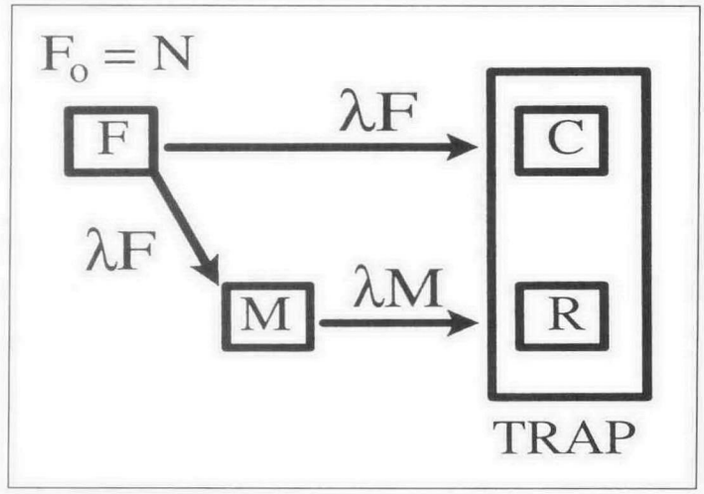

Fig. 1. The original sampling scheme. Traps and marking stations are available simultaneously, and per capita marking rates and trapping rates are both assumed to equal λ. The model projects the number of marked and unmarked insects expected in the traps.

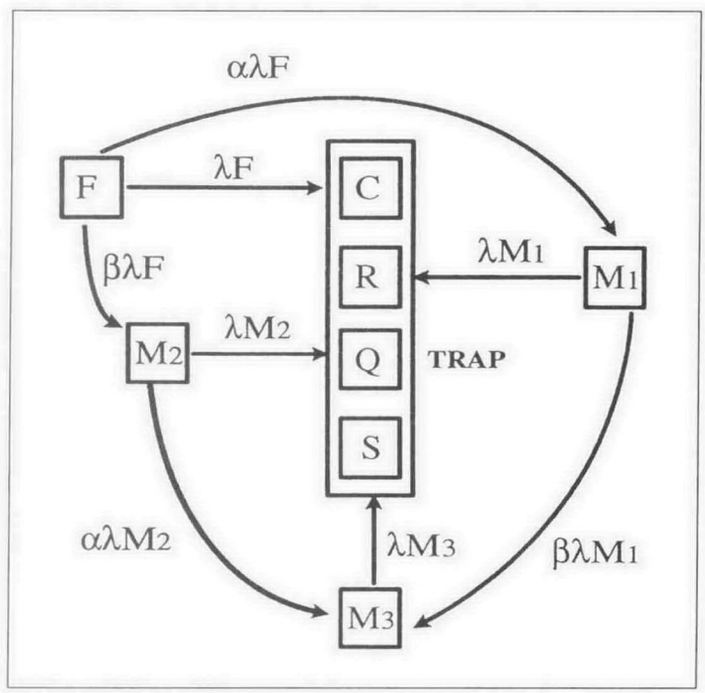

Fig. 2. The double marking scheme. For every trap, there are α red marking stations and β blue marking stations. Insects may arrive in a trap as unmarked, red, blue, or doubly marked. This doubles the degrees of freedom for the model.

and they work well. Generally, the estimate is biased (on average, misses the true value for population size) when the expected number of recaptures is small (Stuart and Ord 1991). Bias is easily corrected using a technique called the jackknife (except for the original 1:1 scheme, which is corrected analytically); the jackknife is a systematic resampling technique which empirically replicates the sampling distribution of the estimate (Efron 1982; Hinkley 1983; Potvin and Roff 1993). The corrected estimate is virtually unbiased unless the expected number of recaptures falls below five (Wileyto, unpublished).

Confidence limits may be obtained either as two standard-error limits, or by using the profile-likelihood technique. The

two standard-error technique uses the variance estimate that comes from the information matrix that is used in the numerical estimation routine (Mood et al. 1974; Lehmann 1983). The profile-likelihood technique uses the likelihood ratio to determine confidence limits (Venzon and Moolgavkar 1988; Arnason et al. 1991). The profile-likelihood approach performed better than the two standard-error approach, producing as short an interval as two standard errors, but including the true value of the population size much closer to 95% of the time.

Adding to the Model

Degrees of freedom may be added to the sampling scheme by borrowing the approach used in the removal method, that of revisiting the trap to make sequential observations. Figure 3 shows a simplified version of the double mark sampling scheme, except that the traps are inspected a second (or third) time. This nets four degrees of freedom for each observation, for totals of eight or twelve. The sequential observations with many degrees of freedom are needed for the next step which is adding other ecological processes to the trapping model. Figure 3 also shows the simplest addition that can be made: recruitment of adults to the free and unmarked class (F). These could come from outside of the population, or from the larval age classes. The recruitment process could contain any process equation that makes sense, and is useful to the modeller. I have begun work with just two models. Model 1 (Fig. 3) works under the assumption that recruitment is constant over the sampling period, and proportional to the adult population by some factor b. This requires the estimation of b, in addition to the usual parameters of initial population size and time. Model 2 (Fig. 3) works under the assumption that recruitment is changing through the sampling period and uses a quadratic function of time to profile that variable recruitment. This requires the estimation of three additional parameters: b_1, b_2, and b_3.

It should be noted here that the trap categories are no longer represented by a multinomial probability model. However, the actual number in each category is represented by a Poisson distribution, where the mean is the expected count based upon solving the differential equations that arise from Figure 3.

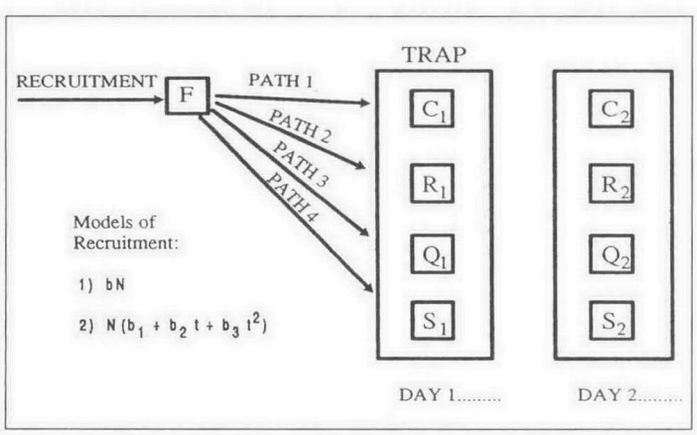

Fig. 3. Another view of the double marking scheme. The paths into the trap are simplified, but degrees of freedom are expanded by having the observer visit the traps more than once. The added degrees of freedom are useful for estimating recruitment to the adult population, the arrow at the left. Thus far, recruitment has been modelled as a constant rate (model 1), or as a quadratic profile over time (model 2).

Performance of the recruitment estimates must also be evaluated by simulation. Parameter b from model A could often be evaluated from a single visit to a trap, although two visits worked more consistently. Confidence limits from a single visit were too broad to be useful. Using two visits, b was biased for small expected numbers of recaptures. Bias adjustment by jackknifing has not yet been tested. Again, the profile-likelihood approach was found to be superior to the two standard-error approach.

Model B has not yet been tested, though I have made trial estimates, using a sampling scheme with three observations. The parameters b_2 and b_3 appear to have very tight confidence limits compared to b_1 or to b from model A.

Applications

I would like to touch briefly upon one of the potential applications for recruitment models. The quadratic function estimated will allow inferences to be made about the age distribution of the underlying larval population. Insects that have colonised a stored commodity often go through a series of boom and bust cycles before the age distribution becomes stable. Although augmented biological control for stored commodities is only now being investigated, we already know that timing of parasitoid release can be very important, especially if only a single release is to be made (L. Smith, pers. comm.). The quadratic model can tell us whether the age classes required by the parasitoid are on an upswing, downswing, or if stationary, whether they are on a crest or in a trough.

Conclusion

Self-marking offers a simple approach to estimating parameters of insect populations. It relies on the construction of dynamic models representing insects arriving at marking stations and traps. Those models might include only changes due to marking and trapping in a closed population, or they may incorporate constructs from ecological models representing change or structure in the population. Applications include inference of population age structure to aid in timing of treatment.

Acknowledgments

Research was conducted as part of the Agricultural Research Service project on Population Ecology and Predictive Models for Packaged Commodities, CRIS No. 6605–43640–018–00D, M.A. Mullen senior investigator. I thank M.A. Mullen for his ideas which set the project in motion, and I thank W.J. Ewens for being my mentor and for introducing me to estimation problems. I would also like to thank R.T. Arbogast, M.A. Mullen, L. Smith, D.K. Weaver for their comments on earlier versions of this manuscript.

References

Andow, D.A. 1984. Estimating relative population density of many species from trap data. Applied Entomology and Zoology 19, 391–392.

Arnason, A.N., Schwarz, C.J. and Gerrard, J.M. 1991. Estimating closed population size and number of marked animals from sighting data. Journal of Wildlife Management 55, 716–730.

Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. (Regional Conference Series in Applied Mathematics). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Hagstrum, D.W., Flinn, P.W., Subramanyam, B., Keever, D.W. and Cuperus, G.W. 1990. Interpretation of trap catch for detection and estimation of stored product insect populations. Journal of the Kansas Entomological Society 63, 500–505.

Hagstrum, D.W., Meagher, R.L. and Smith, L.B. 1988. Sampling statistics and detection or estimation of diverse populations of stored product insects. Environmental Entomology 17, 377–380.

Hinkley, D.V. 1983. Jackknife methods. Encyclopedia of statistical sciences 4, 280–7.

Lehmann, E.L. 1983. Theory of point estimation. (Probability and mathematical statistics series). New York, NY: Wiley.

Mood, A.M., Graybill, F.A., and Boes, D.C. 1974. Introduction to the theory of statistics. (3rd Edition). New York, New York: McGraw-Hill Book Co.

Nichols, J. D. 1992. Capture–recapture models. Bioscience 42, 94–102.

Pollard, J.H. 1977. A handbook of numerical and statistical techniques. Cambridge, U.K., Cambridge University Press.

Potvin, C. and Roff, D.A. 1993. Distribution free and robust statistical methods: viable alternatives to parametric statistics? Ecology 74, 1617–1628.

Seber, G.A.F. 1982. The estimation of animal abundance. (Second Edition). New York, NY, MacMillan.

Shore, T.L., and McLean, J.A. 1988. The use of mark recapture to evaluate a pheromone based mass trapping program for ambrosia beetles in a sawmill. Canadian Journal of Forest Research 18, 1113–1117.

Stuart, A. and Ord, J.K. 1991. Kendall's advanced theory of statistics. (5th edition). Volume 2. New York, NY: Oxford University Press.

Venzon, D.J., and Moolgavkar, S.H. 1988. A method for computing profile-likelihood based confidence intervals. Applied Statistics 37, 87–94.

White, G.C., Anderson, D.R., Burnham, K.P. and Otis, D.L. 1982. Capture-recapture and removal methods for sampling closed populations. Springfield, VA: National Technical Information Service, U.S. Government Printing Office.

Wileyto, E. P., Ewens, W. J. and Mullen, M. A. Markov/recapture population estimates: a tool for improving interpretation of trapping experiments. Ecology, in press.