

# CHAPTER 9

## THE FUTURE

Until recently, most biologists were isolated from quantitative methods and computing devices, but they now are beginning to be trained in basic statistics, deterministic models, simulation, and the use of digital computers. Scientists in the biological fields are progressing rapidly into sophisticated fields of analysis such as stochastic processes, multivariate analysis, and nonlinear and dynamic programming. We see this as a positive trend.

### Biological Developments

Major biological developments in the future probably will involve improvements in field techniques, to reduce the variability in capture probabilities. As our knowledge about the mechanisms that cause differences in capture probabilities increases, experiment design can be improved to eliminate the variability. Biotelemetry provides one means of studying the reasons why capture probabilities vary between animals, and the ways in which animals react to the placement of traps in their home ranges. For example, does the strangeness of a trap placed in the middle of an animal's home range tend to scare the animal away? Does its curiosity make the same animal investigate a trap placed in a strange area outside of its home range? Where in an animal's home range should a trap be placed to maximize the chance of catching that individual? Such questions will be answered as we learn more about the behavior of animals in relation to capture-recapture experiments.

Among the possibilities for eliminating some of the variability of capture probabilities are better ways to mark animals and later to differentiate marked from unmarked animals. One promising prospect involves methods in which marked and unmarked animals are distinguished without actually being recaptured. Possibilities include the use of electronic tags and an electric eye gate: one counter would tally all animals that pass through the gate and, in addition, the electronic tag would cause a second counter to tally only tagged animals. A number of animals would be tagged initially to introduce a marked segment into the population. The actual experiment to determine the proportion of marked animals would not increase the number of marks, but would only make counts of marked versus unmarked animals (*Rice and Harder 1977*). Because the animal would not be captured, behavioral response to trapping would be eliminated during the study's second stage. Moving the counting devices would help reduce heterogeneity of capture probabilities. Aerial surveys could be used to count animals such as mule deer, and an electronic device could be used to determine which animals carry the electronic tag. Such a scheme would eliminate the observer bias of spotting brightly colored neck collars before the animal is seen, or of failing to see a dull colored collar on a marked animal and thus counting it as unmarked.

A conceptually similar idea is to inject animals with a chemical, which is eliminated in the feces. Feces are collected, and the number of marked and unmarked feces are used to estimate the proportion of marked animals in the population. *Pelton and Marcum (1975)* used the method to estimate the number of black bears in the Great Smoky Mountains National Park, and *Davison (1980)* used it to estimate coyote populations in the Idaho National Environmental Research Park. In both studies, the radioactive materials used to tag the animals made detection in the feces relatively easy with radiation counters. An extension of the same idea could use stable chemical elements that are rare in the environment, but for which analytical techniques exist, to determine if the amount of material in the feces was above the

background for the element and thus from a marked animal. Fecal-tagging methods avoid the problem of behavioral response to trapping, unless the injected substance causes a change in defecation rates. However, differences in defecation rates between individuals or age and sex classes may cause significant heterogeneity in defecation rates and hence in the probability of finding the animals' feces.

The old adage about developing a "better mousetrap" is true. Improvements in all types of gear are needed, including increased overall reliability, the ability to catch several animals in a single trap, reliability in freezing temperatures, prevention of trap deaths or escape of captured animals, and reduced costs per unit. Often, the tags now being used fall off, or become unreadable, or cause mortality or infection. The use of microtaggants is one example of needed innovations (see *Johns 1979*).

Perhaps the least explored subject deals with alternative trapping designs. At present, only three general designs are widely used: (a) traps in a single line—a design totally without merit, even for use as an index; (b) traps in a rectangular or irregular grid; and (c) a grid with assessment lines. Other designs might be possible and could even facilitate the direct estimation of density.

Finally, there is a critical need for more studies like those conducted by *Mahon (1980)*, in which the true size of the animal population is known, and the capture-recapture or removal sampling is conducted to estimate  $N$ . This type of study enables a strict comparison of  $N$  versus  $\hat{N}$  and is most meaningful if it is replicated. Studies like those conducted by *Edwards and Eberhardt (1967)* and *Carothers (1973b)* are very informative, but they allow comparison and performance of estimators based on only one outcome of the stochastic sampling process.

We are extending a plea to authors and editors of manuscripts involving the use of capture-recapture methods to document any experiment features that might help others to design experiments with the same species. Most commonly, only the population estimate and its standard error are reported, with no reason given for using the particular estimation technique and no estimates of capture probabilities provided. An increase in the amount of such information would facilitate the construction of efficient and successful trapping designs. For example, when the results of several trapping experiments involving house mice and using the program CAPTURE to analyze the data are reported, other researchers planning similar work will want to know which models tended to be selected for estimation and the estimated probability of capture. In a sense, such information is as valuable to other experimentors as the actual population estimates.

## Statistical Developments

Statisticians have contributed to capture-recapture theory on two levels: first, in the development of new models and their associated methods of analysis, and second, in the evaluation of established methods (for example, investigations of the robustness and efficiency of estimation methods and evaluation of the power of tests of assumptions). More work needs to be done in both areas. Meaningful contributions on new, realistic models probably will be made by statisticians working with biologists, as opposed to mathematically trained statisticians with no background in biology, working alone. There is a large potential for improved capture-recapture studies if biologists and statisticians work together to design capture and removal studies and associated models (cf. comments of *Hartley 1980:2*).

We foresee the continuing development of both realistic new models and special cases of existing models. Development of special cases will be motivated by the real problems biologists face. Development of new, realistic models will be based on experience with the capture-recapture method and the insights to be gained by extensive analysis of existing data. In both regards, the joint efforts of biologists and statisticians are needed. For such cooperation to work well, the statistician needs to know something about biology, including the realities of field work, and the biologist needs to have a basic training in statistical methods, preferably including knowledge of and experience with computers. Given teams with experience in obtaining and analyzing capture-recapture data, we foresee the development both of models incorporating specific biological characteristics of target populations and of optimal statistical analyses.

When statisticians and biologists work alone, the statistician often builds elegant but unrealistic models, while the biologist, who knows the realism the models should have, cannot formulate efficient statistical analyses.

**Closed Models.** A difficulty plaguing both open and closed models, but better understood in closed models, is the poor confidence interval coverage on population size  $N$ . The difficulty derives from two problems: (1) the normal distribution is a poor approximation to the sampling distribution of  $\hat{N}$  for many models, even for moderate sample sizes; and (2) there is a very high sampling correlation between  $\hat{N}$  and  $\hat{se}(\hat{N})$ . The latter problem occurs because in most models  $\text{var}(\hat{N})$  generally equals  $N$  times some function of the capture probabilities and  $t$  (capture occasions), and therefore the estimated standard error of  $\hat{N}$  is proportional to  $\sqrt{\hat{N}}$ . This correlation causes  $\hat{se}(\hat{N})$  to be overestimated when  $\hat{N}$  is larger than  $N$  (in this case coverage exceeds 95%) and to be underestimated when  $\hat{N}$  is less than  $N$ . Of course, in practice we do not know  $N$ , hence we do not know which situation prevails, nor its seriousness. Values of  $\hat{N}$  much smaller than  $N$  generally result in a 95% confidence interval that does not cover  $N$ , whereas values of  $\hat{N}$  much larger than  $N$  result in an interval that almost always covers  $N$ .

The problem of nonnormality may be solved by determining better approximations to the sampling distribution of  $\hat{N}$ , directly or through transformations that improve the normality assumption. For the jackknife estimator  $\hat{N}_h$  of Model  $M_h$ , the estimated variance typically underestimates the true sampling variance, although the normality assumption on  $\hat{N}_h$  seems acceptable. In this situation, a bootstrap method (Effron 1979) might be used to advantage to estimate the sampling variance of  $\hat{N}_h$ . In other models, extensive simulation may provide insights into the distribution of  $\hat{N}$ , so that more reliable confidence intervals can be constructed; the concept is related to the idea of using the bootstrap method based on a given set of data to investigate the distribution of  $\hat{N}$ .

Whether using the uniform minimum variance unbiased estimator (UMVUE) of  $\text{var}(\hat{N})$  will solve the correlation aspect of the coverage problem is not clear; however, developing these variance estimators is worthwhile. Berg (1976) has developed the UMVUE for  $\hat{N}$  of Model  $M_1$ . Similar work remains to be done for the removal estimators and Model  $M_0$ .

In general, asymptotic results for the estimators  $\hat{N}$  and  $\hat{se}(\hat{N})$  and for tests of assumptions should be avoided as much as possible. Such results can be especially poor for capture-recapture models. Alternative estimators to the ML approach have not been investigated widely for capture-recapture data. For example, minimum chi-square, minimax, and decision theoretic or information theoretic approaches may have merit. Calculations for all of them are highly computer oriented.

Bayesian statistical methods have found little favor in wildlife studies, partly because there has been almost no investigation of Bayesian methods for use with capture-recapture studies. We doubt they will find much favor, except in modified form. For example, specific types of probability distributions for the capture probabilities in the heterogeneity model may be appropriate. If the distributions could be deduced on the basis of the analysis of capture-recapture data, they might lead to improved estimators of population size under Model  $M_h$ . The same idea applies to the generalized removal model; if the heterogeneity could be modeled in some parametric form, improved removal estimation might result. Finally, the model selection rule of Otis *et al.* (1978) could be improved. As now formulated, it unrealistically assumes equal prior probabilities on all eight models. Unfortunately, we do not yet know on a species-by-species basis what better priors to use. An improved discriminant function rule should include the option of unequal prior distributions on the models in conjunction with specific information on which priors are appropriate for use in specific situations.

There are other ways to improve the model selection rule. One quick improvement would be to allow unequal covariance matrices in the discriminant function analysis by using quadratic discriminant functions. The quadratic equations could then replace the linear equations now available in program CAPTURE. There is also a problem with the coherence of the selection criteria with respect to Model  $M_1$ . Often, when the selection criterion is 1 for Model  $M_{th}$ ,  $M_{tb}$ , or  $M_{tbb}$ , it is zero for Model  $M_1$ , but

substantially more than zero for Model  $M_o$ , indicating that the time function is a significant variable, along with others, but that Model  $M_o$  is preferred over Model  $M_t$  as an explanation of the data. We suspect that this incoherence relates to problems with the goodness of fit test for Model  $M_t$ , which requires a large amount of data and apparently has very low power. Conceptually, a selection rule that recognizes the relations among the models might help solve this "coherence" problem. In the extreme, a nonparametric discriminant function approach may be possible.

The power of the tests is a general problem; none appear to have very high power. Thus, the search for better tests of assumptions is a fruitful area for future research in capture-recapture methods. What approaches might be tried are not clear. Randomization methods are worth considering for these tests, as is the related bootstrap method of *Efron (1979)*.

There is a clear need for estimation procedures for Models  $M_{th}$ ,  $M_{tb}$ , and  $M_{tbb}$ . In each case, the parameters in these unrestricted models cannot be estimated individually. Thus, progress requires some way to reduce the number of parameters. In Model  $M_{tb}$ , the assumption that first-capture and subsequent recapture probabilities are proportionally related across the time effect would make  $N$  estimable. The reality of this assumption could be explored. A jackknife or bootstrap method might work for Model  $M_{th}$ . Alternatively, the introduction of moderately specific distributions on the capture probabilities for the heterogeneity component would reduce the number of parameters to a tractable number. We hold little hope for estimation under Model  $M_{tbb}$ .

Models intermediate to those in *Otis et al. (1978)*, such as models to allow recaptures to bear on the estimation of  $N$  in Models  $M_b$  and  $M_{bb}$ , also may be useful. To obtain such models, however, we must postulate a relation between first and subsequent capture probabilities.

The log-linear model approach (*Cormack 1981*) deserves attention; its properties for analysis of capture-recapture data should be determined. Whether it works well for real capture-recapture data is not yet known, and one good simulation study would go a long way toward answering this question. The study should generate the data by using models based on different capture probability structures. The log-linear theory has no obvious relation to any capture-recapture probability structure except for Models  $M_t$  and  $M_o$ . When the log-linear method of analysis selects a model with interaction terms, it is difficult to relate the model to heterogeneity or behavior (or both) or to some interaction among the sources of variation in capture probabilities. This difficulty will hamper biological understanding of capture results and hence hamper the design of future studies and the development of better methods of analysis. More research ought to lead to a better understanding of the results of log-linear analysis.

Log-linear models also may be improved by recognizing the time-ordering of captures. Because the time-ordering of the capture occasions is not taken into account in present log-linear models, a proper analysis of Models  $M_b$  and  $M_{bb}$ , the models requiring removal estimators, is especially difficult. True removal studies clearly do not fit into the log-linear framework. All these considerations lead us to believe that the log-linear approach may not solve the problem of developing a unified approach to analysis of capture data.

Relatively fewer areas of research are left for removal models, compared with capture-recapture models. All major research on mathematical aspects of the constant capture probability removal model has been done. There are ways to continue to investigate and modify this model (*Skalski and Robson 1979, Carle and Strube 1978*), but they are minor changes or refinements to the basic model and the ML estimator. However, investigations and improvements in the area of removal studies with heterogeneous capture probabilities are needed. The generalized removal estimator is the only existing treatment of this situation. Alternative approaches could model the degree of heterogeneity or could empirically model a relation between the conditional average capture probabilities on each occasion ( $\bar{p}_j$ ). For example, perhaps a model such as  $\bar{p}_j = a/(a + bj)$ ,  $\bar{p}_j = a/(a + b + j)$ , or  $\bar{p}_j = p + e^{-bj}$  would lead to more robust estimation for removal studies.

A final subject of concern is estimation of density. Improvements in the method presented in Chapter 5 and in *Otis et al. (1978:67-74)* are surely possible. The quite different approach of using assessment lines has not received enough attention to determine its validity and usefulness. A combination of statistical

modeling and empirical research on assessment lines seems appropriate. A method, based on trapping, for the direct estimation of density without first estimating  $N$ , then estimating the area, would be very attractive, if one could be found.

**Open Models.** This primer deals with robust analysis of capture data from closed populations. However, there is a tremendous need, and hence a challenge, to develop robust methods of estimating population size for the open models discussed briefly in Chapter 8. The number of possible open models is so large that a comprehensive approach would be very difficult. To simplify this discussion we will consider the Jolly-Seber model (*Jolly 1965; Seber 1965*) as a baseline from which to proceed with new methods.

For one age class and time-only variation, the Jolly-Seber model provides an adequate approach (Chapter 8). However, there is a need for both more restricted and more general models. More restricted models are obtained by reducing the number of parameters; the reduction tends to reduce the realism of the model, but may achieve an overall gain in effective estimation of the key parameters. In particular, the survival parameters can be reduced from  $t - 2$  to 1 by assuming a constant survival rate. Similarly, the recruitment parameters may be reduced by assuming a constant number of recruits per trapping period, assuming a relation between recruitment rate and population size, or otherwise modeling a relation into the recruitment parameters to reduce their numbers. Both approaches to developing more restricted Jolly-Seber models are being pursued.

The need for generalized open capture-recapture models to allow for other sources of variation in parameters is at least as great as the need for more restricted models. The generalized models will require more and/or different parameters to gain greater realism, hence robustness, with respect to the effects of behavior and heterogeneity. To date, there has been no incorporation of heterogeneity, such as in Model  $M_h$ , into open models. We believe the failing is very serious, because both the literature and our own experience with closed models strongly suggest that heterogeneity is always present to some degree. When it is a strong source of variation in capture probabilities, it will seriously bias estimation of population size. The only attempts to allow for heterogeneity that we are aware of have been attempts to develop age-stratified Jolly-Seber models, analagous to those of *Brownie et al. (1978)* for banding data. *Pollock (1981a)* has developed such age-stratified open models. Age, however, is only one potential source of heterogeneity; many sources are not identifiable and hence cannot be overcome by stratification.

Within the next decade we expect to see many developments of specialized and generalized Jolly-Seber models. The worth of these models must be evaluated by analytic and simulation methods and by application to real data. Application to real data, of special importance, will require prompt implementation of these methods in user-oriented computer programs. The state-of-the-art computer program for dealing with open models is POPAN-2 (*Arnason and Baniuk 1978, 1980*), which has many useful features. It implements all the Jolly-Seber models and many of Arnason's innovations, but it implements few of the special cases achieved by assuming restrictions on the survival, recruitment, or capture probability parameters. Nor has any of the work of *Pollock (1975)* been implemented yet. These comments are not a criticism of POPAN-2; rather, they are suggestions for future developments.

We recommend the extensive analysis, or re-analysis, of good capture-recapture data for open-model studies, using program POPAN-2. Publication of the results will help resolve the question about which of the existing models are useful with real data.

The problem of density estimation ( $D = N/A$ ) for open models also must be addressed. There is virtually nothing in the literature about this problem.

To summarize, a great deal of work remains to be done on open models. Research on most of the problems is ongoing, and we expect to see substantial progress in the next 5 years. The challenge will be to evaluate the methods quickly and to make them available for routine applied use. Certainly by the end of this decade there should be a complete synthesis of reliable, efficient methods for the analysis of

capture data assuming an open model. Although no such synthesis exists now and one would be premature in the next 5 years, it will be critical to further progress.

## Computer Programs

Computer programs have become essential in the analysis of capture-recapture and removal studies. The iterative nature of many estimators of population size under closure makes them nearly impossible to compute with a hand calculator, and the testing and model selection procedures are tedious. The notation and algebra for the estimators of open-population parameters are difficult, and recently developed models do not have closed-form estimators. In all cases, rounding errors, especially for the estimates of sampling variances, can be serious on a calculator. Now and in the future, a comprehensive analysis of any set of multiple capture data will require the use of sophisticated computer programs.

In the next decade, we foresee the development of a comprehensive, integrated series of computer algorithms, similar in concept to the SAS (Statistical Analysis System) program (SAS 1979). The program should be in FORTRAN and should be machine independent, at least for common makes of computers in the medium-to-large size class.

The program should include a data maintenance capability that is both sophisticated and easy to use, à la Arnason and Baniuk (1978). The basic data must be accessible to any of the analysis, testing, model selection, or display routines. The system should operate in an interactive, remote-batch, or batch mode at the option of the user.

An extended graphics display routine will be advantageous. The routine should allow, for example, display of a three-dimensional histogram showing the number of animals caught per trap on the grid or subgrid, simple histograms of the data by age, sex, and weight, and figures showing estimates, standard errors, and confidence intervals for the results. A versatile routine to allow students or researchers to simulate data from fixed parameters is a powerful feature of such programs (Arnason and Baniuk 1978).

We expect the number of separate algorithms in the system to increase in a manner similar to the way additions and improvements have been made to the BMD series (Dixon and Brown 1979). Tables 9.1, 9.2, and 9.3 summarize some types of studies from which data are available, the analysis methods used, and the existing programs. Possible additions and improvements are discussed below.

Many other methods could be added, including the many catch-effort (Pella and Tomlinson 1969) and change-in-ratio methods (Paulik and Robson 1969; Otis 1980). Several ML estimators have regression-type approximations that might be added if they were more robust. Finally, many less general methods might find a place in a comprehensive computer system; estimators for instantaneous rates are an example.

A good computer system and its accompanying software are now mandatory, if biologists are to benefit from the statistical and theoretical advances made in the past decade. A general, flexible, easy-to-use system will be a great help in future research on biological populations. A prediction made in 1969 by Overton and Davis (1969:404) reinforces our argument.

Computers will soon prove of very great value in the routine processing of census and survey data. When they become generally available, it will be desirable to advance to even more realistic and complex solutions to the problems; there will be no premium on simplicity, so long as the users understand the principles and are able to comprehend the constraints and limitations of the models on which the computer solutions are based.

**TABLE 9.1.** Analysis methods for studies where data are available on only marked members at time j.

Method	Existing Program, Reference
Seber (1970); Robson and Youngs (1971)	ESTIMATE, Brownie et al. (1978)
Other models allowing age-independent parameters	ESTIMATE, Brownie et al. (1978)
Brownie and Robson (1976)	BROWNIE, Brownie et al. (1978)
Other models allowing age-dependent parameters	BROWNIE, Brownie et al. (1978)
Johnson (1974)	--- Johnson (1974)
Seber (1971); Cormack in Fordham (1970)	--- Fordham (1970)

**TABLE 9.2.** Analysis methods for studies where data are available on only unmarked members at time j.

Method	Existing Program, Reference
Zippin (1956, 1958); Moran (1952)	CAPTURE, Otis et al. (1978)
Generalized Removal Method; Otis et al. (1978)	CAPTURE, Otis et al. (1978)
Chapman and Robson (1960); Robson and Chapman (1961)	--- ---
Dupont (1976)	FISH, Dupont (1976)

**TABLE 9.3.** Analysis methods for studies where data are available on both marked and unmarked members at time j.

Method	Existing Program, Reference
<b>Closed Models</b>	
Darroch (1958, 1959); Burnham & Overton (1979);	
Pollock (1974)	CAPTURE, Otis et al. (1978)
Other estimators assuming closure	CAPTURE, Otis et al. (1978)
<b>Open models</b>	
Arnason (1972, 1973)	---
Cormack (1964)	---
Darroch (1961)	---
Fisher and Ford (1947)	POPAN-2, Arnason & Baniuk (1978, 1980)
Jolly (1965); Seber (1965)	POPAN-2, Arnason & Baniuk (1978)
Jolly (1979)	--- Bishop & Shepard (1973)
Manly and Parr (1968)	---
Pollock (1975); Robson (1969)	---
Pollock (1981a)	---
Seber (1962)	---