Routledge
Taylor & Francis Group

Capture-recapture: before and after EURING 2000

GEORGE A . F. SEBER[1] **& CARL J. SCHWARZ**[2], [1]*Department of Statistics, University of Auckland, New Zealand and* [2]*Department of Statistics and Actuarial Sciences, Simon Fraser University, Burnaby, Canada*

ABSTRACT   *Capture-recapture studies and analyses have become an important tool for the study of bird populations. One reason for the rapid advancement in this area has been the EURING conferences where population biologists and statisticians meet to review recent progress, identify areas that require further work, and work collaboratively to solve real world problems. In this paper, we forecast the needs for future research in this area and review the recent conference to try and identify what questions are yet unsolved.*

This EURING conference was dedicated to Dr George Seber who was the author of a number of key papers and whose name is synonymous with 'The estimation of animal abundance and related parameter' (Seber, 1982). He has retired from working in this field.

1 Preface

The EURING meetings are the premiere conference for advances in capture-recapture methodology, particularly as applied to studying bird populations. The meeting typically alternates between Europe and North America. EURING 2000 was held in early October just outside the Pt. Reyes Bird Laboratory in California.

This paper will begin with a forecast of where capture-recapture may head over the next few years followed by a summary of the key conference ideas and challenges for the future.

2 Where is capture-recapture going?

It is interesting looking back over about 50 years of papers on animal abundance to see how far we have come. Computers have revolutionized our subject as they

Correspondence: C. J. Schwarz, Department of Statistics and Actuarial Sciences, Simon Fraser University, Burnaby, British Columbia, V5A 156 Canada. E-mail: cschwarz@stat.sfu.ca

have done for the whole of statistics. In the early days, the aim was to set up capture-recapture models that had explicit maximum likelihood estimates rather than to investigate more realistic models possibly requiring iterative solution. This has, of course, all changed. Models are now routinely fitted with many hundreds of parameters, and numerical solutions rather than analytical formulae are now the order of the day in fitting the models and determining goodness-of-fit.

Advances in the field have been numerous and the growth in the basic literature explosive. For example, Seber (1973) had over 400 references; in Seber (1982) this had grown to over 1000 references. And in subsequent reviews (Seber, 1986, 1992; Schwarz & Seber, 1999) the number of new papers increased to over 350, 500, and 400 respectively, with the latter number now excluding all application papers and consisting of just new theoretical advances!

2.1 Closed population—a plethora of methods. Consolidation required

Closed population models are the 'simplest' of capture-recapture studies and form one of the basic components of the robust design. However, there are many different methods to estimate the abundance in these populations. These include the standard maximum likelihood methods outlined in Otis *et al.* (1978); the jackknife models of Otis *et al.* (1978); the coverage models of Chao, and colleagues (Chao *et al.*, 1992); the martingale estimating equation models of Becker(1984), Yip (1989), Lloyd & Yip (1991) and others; the log-linear models of Cormack (1989); the logistic models of Huggins (1989) and Alho (1990); the non-parametric MLE methods of Norris & Pollock (1996); the mixture models of Pledger (2000); the log-linear constrained models of Evans *et al.* (1994); and Bayesian models (Underhill, 1990). The driving force behind many of these techniques has been heterogeneity in capture probabilities among individual animals; a common feature of many animal populations. Some of the models incorporate heterogeneity by using covariates or stratification, an approach to modelling that is being used more and more.

With this plethora of methods, the practitioner is left in a quandary—what methods are appropriate for what conditions and types of studies. What is needed here is a comparison of the various closed models with respect to both efficiency and robustness. Also, further research is needed on interval estimation. Typically, intervals based on profile likelihoods are more reliable than those based on the (hopefully) asymptotic normality of the maximum likelihood estimates. However, in the case of non-maximum likelihood methods, there are no obvious suggestions on how to improve their performance.

2.2 Open population models—the Cormack-Jolly-Seber model

In Cormack-Jolly-Seber studies (Cormack, 1964; Jolly, 1965; Seber, 1965), marked animals are released and followed over time. Emphasis is on estimating survival; abundance cannot be estimated as the process by which marked animals are obtained is not modelled, so that all inference is conditional upon the release. The key paper by Lebreton *et al.* (1992) started a revolution that is still ongoing. The adoption of an ANOVA type framework and model selection methods are the core to most of the modern methodology.

Live-recaptures, resightings and dead-recoveries in tandem with both age-dependence and tagging effects can be incorporated into a single model (Barker,

1999). Freeman & Morgan (1992) and subsequent papers have developed elaborate models to deal with birds ringed as nestlings. It is now not uncommon to see models with several hundreds of parameters. However, this raises the twin issues of the identifiability of parameters, and the proliferation of parameters (cf. Burnham *et al.*, 1995). Such elaborate models, have taken us about as far as we can probably go in terms of numbers of parameters. There is only a limited amount of data available so that as the number of parameters goes up, the efficiency of estimation goes down. In addition, sparse data may limit the number of parameters that can be fitted; often unrealistic restrictions need to be applied to the parameters to achieve estimation.

Clearly this trend in increasing complexity needs to be reversed in some way towards models with fewer parameters. Perhaps what may be needed are regression models where much of the variation and overdispersion can be incorporated into the 'error' term. For example, if parameters are interpreted as mean values (e.g. the average probability of capture), heterogeneity can often be incorporated into a multinomial distribution. However, regression models may do it more effectively by including any variation about the mean in the error term. For example, suppose $X_i = Np_i + \varepsilon_i$ as in a Binomial model, and $p_i = p + v_i$, then $X_i = Np + Nv_i + \varepsilon_i = Np + n_i$. Because of all the various sorts of variation taking place in a study, we may not need to be too concerned about non-homogeneous errors or non-normality of the errors. A good example where the error term may 'hide a multitude of sins' is Leslie's (unweighted) catch-effort regression model (Seber, 1982, Section 7.1.3). The same idea can be applied in the framework of a generalized linear model, but with a different error structure (e.g. normal rather than Poisson generated). Regression models also allow for the ready incorporation of covariates. Of particular interest is Burnham's random effects model (Burnham, 2001; Burnham & White, 2002) in which a long sequence of survival estimators can be summarized in terms of two parameters, a mean and a variance. We envisage mixed models with some random effects will be used more widely. Meta-analysis, used in subjects such as education, definitely has a place here as another way of combining information and reducing the number of parameters. Radio tagging also has a niche, although handling the large numbers of radio tags needed for efficient estimation creates cost and logistical problems.

Linked with the proliferation of parameters is the related problem of model selection, especially when there are several close contenders, all equally useful and biologically plausible. Likelihood ratio tests have been traditionally used to discriminate between models. Because it was computationally difficult to fit models, alternatives to the usual likelihood ratio test were used, such as score tests (e.g. Catchpole & Morgan, 1996) where only estimates under the null hypothesis, or Wald tests where only estimates under the most general model, are required. Techniques for variable selection analogous to those used in regression situations (e.g. forward, backward, and stepwise selection) have been proposed but these suffer from the same problems as found in regression settings. All of these hypothesis testing methods only address one side of the problem—if a factor is 'significant', then one wants to include it in the model, but the 'cost' in terms of lost precision is not addressed.

For these reasons, there has been a dramatic shift away from using hypothesis tests in model selection to use measures such as AIC to compare competing models. Researchers are becoming more aware of the dangers associated with selecting a model and then using it to estimate parameters without taking into

account the process of model selection (Burnham & Anderson, 1998). As a result, model selection is now somewhat of an industry. Basing a final estimate on some sort of (weighted) average of estimates from related models is clearly a reasonable thing to do, although care is needed when there are 'good' models giving very different estimates. One can use either an ad hoc non-Bayesian approach for choosing the weights, as in Buckland *et al.* (1997), or a Bayesian approach using Bayes' factors, as in Hoeting *et al.* (1999). More research will no doubt continue in this field both with regard to the choice of selection criteria (which we need to know more about) and the selection of the weights.

The CJS models can also be used to estimate movement of the animals. The models were originally developed to investigate physical movement among different areas (Schwarz *et al.*, 1993; Brownie *et al.*, 1993) but 'movement' can be generalized to any movement among states. For example, Nichols *et al.* (1992) estimate movements among weight classes. Lebreton *et al.* (1999) and Lebreton & Pradel (2001) showed how the multistrata approach can be used to investigate more general questions in ecology, or to combine sources of information. Similar comments about parameter proliferation also apply here, but the danger is greater as the number of parameters increases with the square of the number of states. Again, some sort of model simplification is needed—perhaps the movement rates can be modelled as functions of distance between the sites or some underlying smooth curve that is discretized (e.g. Schwarz & Dempson, 1994).

Many CJS studies are observational rather than manipulative. It was therefore good to see a session labelled 'Formal experiments with ringed birds'. In practice, experimental design tends to be more of an art than a science, but the formal aspects of design do help us to identify some crucial issues. These include such things as choosing the blocking factors; determining the nature, size and numbers of the experimental units; and deciding which effects are fixed and which are random (Schwarz, 2002). This is a largely untapped area in the CJS model where, in the past, all effects are treated as fixed and there is a single residual error term.

2.3 Open population models—the Jolly-Seber model

Jolly-Seber (JS) studies are similar to those of CJS studies except that now the process by which new animals are captured, marked, and returned to the population must be modelled. Ironically, the first applications of mark-recapture methods in the early 1960s was to estimate abundance. Survival was a nuisance parameter that had to be estimated in order to obtain estimates of the abundances at each sample occasion. Because of the close similarity between these two types of studies, the concerns and research directions raised for the CJS models are also applicable to the JS models.

The primary weakness in the JS models is the severe effect of heterogeneity in capture probabilities upon the estimates of abundance. Consequently, some of the approaches used in closed populations to deal with the problem have now been extended to open populations, and this process will continue in the future. For example, the coverage method continues to be a fruitful source of model developments. The paper by Hwang & Chao (1995), which estimates the biases of the Jolly-Seber estimates due to the effect of unequal catchabilities, is an important extension of the coverage method into the domain of open populations. Pledger & Efford (1998) developed two additional methods to correct this bias. They derived formulae for the asymptotic relative bias as a function of the coefficient of variation

in the capture-probabilities (which is estimated directly from the data) or used a simulation-inverse- prediction correction.

Population abundance is also related to another session on the estimation of λ (population growth), a parameter which, along with applications of the Leslie matrix, is regarded as an important part of wildlife management. However, the use of λ is problematical as it requires a stable age distribution, something not easy to prove. In addition, estimating population growth from the CJS model and population abundance from the JS model must be closely related. It is not clear from the current literature the equivalence between these two approaches.

2.4 Robust design

On a number of occasions we have emphasized the importance of Pollock's (1982) robust design. The robust design combines the features of closed populations during the secondary sampling occasions with the features of open population across the primary periods. It was originally developed to reduce the effects of heterogeneity upon the estimates of abundance but has now been extended to differentiate between *in situ* growth and immigration (Nichols & Pollock, 1990). Schwarz & Stobo (1997), Kendall & Nichols (1995) and Lindberg *et al.* (2001) were able to separate temporary and permanent emigration. There is clearly great potential here for developing a whole range of models that can effectively separate out the various processes involved in an open model. The robust design has an even greater flexibility as it is essentially a combination of two different modelling processes. A more general study design can therefore be used during the secondary periods, e.g. Gould & Pollock (1997) used a catch effort study during the secondary periods. However, with any such structured design, the question always arises about the best allocation of resources among the various levels of the design.

2.5 General methodological issues

2.5.1 Bayesian methods.   In the early days, Bayesian models for closed populations occasionally appeared, but never really took hold because of unrealistic priors and the computational difficulties in calculating posterior distributions. In addition, Bayes estimators tended to be very sensitive to the priors used (see for example Garthwaite *et al.*, 1995). Freeman (1990), in an unpublished PhD thesis, suggested applying Bayesian methods to the open model based on recoveries from birds banded as nestlings, but ran into the previously mentioned computational difficulties. A non-Bayesian solution was given by Freeman & Morgan (1992). Markov chain Monte Carlo methods and Gibbs sampling then came on the scene providing the tools for handling most sensible priors. The method avoids the calculation of the complex integrals needed for completely specifying posterior distributions, and it readily leads to a sensitivity analysis for the chosen priors. Two papers on open models then appeared by Dupuis (1995) and Vounatsou & Smith (1995). Dupuis developed a Bayesian framework for the Arnason-Schwarz stratified open model with capture- recapture data. He utilized the idea of a missing data structure to develop the Gibbs sampling algorithm, a technique that has promise. However, the priors were assumed known and the sensitivity of the estimators to the priors was not investigated. Vounatsou & Smith (1995) further developed Freeman's Bayesian model and compared their estimates with the maximum likelihood estimates of Freeman & Morgan (1992). The how-to-do it papers by Brooks *et al.* (2000a,

2000b, 2002), and which spells out general Bayesian methods for recovery and recapture studies is a welcome addition to the literature.

Brooks *et al.* (2000a) also raised the spectre of sensitivity to the prior. It is clear that priors should be chosen on the basis of biological information and not mathematical convenience, and a sensitivity analysis should always be carried out. The role of Bayesian methods still needs further investigation when dealing with sparse data, non-identifiability and non-estimability of parameters, and out-of-range estimates. Finally, while the computational issues that inhibited the use of Bayesian methods in the past have diminished, the development of a standardized computer package for capture-recapture studies is still required before these methods will be more widely used.

2.5.2 Covariates.   We have already mentioned the possibility of incorporating covariates into models. Effort data has been incorporated into fisheries models for some time. However, one of the first papers to incorporate environmental variables as covariates in the animal population scene appears to be North & Morgan (1979), and the theory was spelled in the key papers by Pollock *et al.* (1984) and Pollock (2002). Since then, covariates have been used extensively, with logistic models in particular being useful for modelling probabilities (e.g. Lebreton *et al.*, 1992). With improved monitoring devices we can expect covariate activity to increase, not only as a means of allowing a greater variability of parameters (e.g. some parameters do not need to be made equal for identifiability), but also as a means of estimating the standard parameters more efficiently by the injection of more information. However, we feel that covariate relationships are not always studied as carefully as one might hope before they are used.

2.5.3 Model diagnostics.   Diagnostics to determine model adequacy are often given minimal attention when fitting population models. Various diagnostics involving residuals are available for linear and log-linear models (e.g. Lee & Seber, 2001), and we need similar tools for our models and for studying covariate relationships. Chi-squared goodness of fit tests can have notoriously low power, particularly when a large number of parameters are fitted. A closer look at the data may tell us a bit more about what is going on. Cormack (1989) made some suggestions about appropriate residual plots, but the key problem in capture-recapture models is the sparseness of the individual capture histories. Some sort of automated collapsing of histories to view slices of the data will be important, such as plotting the residual versus the number of capture occasions.

2.5.4 Planning.   Finally, there is one area where much more work needs to be done. Suppose someone came to you to ask for advice on how to design a capture-recapture study for a closed population. How would you go about it? What sample sizes are required, and how many recapture samples are needed? Unfortunately, apart from a few books and papers, little attention is often given to this problem.

3 EURING responds to the challenge

The EURING 2000 meetings were loosely organized around seven major sessions. The proceedings of this conference (this issue) contain the keynote and other papers. What follows are impressions about the key points raised in each session and how the sessions address some of the future research directions raised above.

3.1 Evolutionary biology

Any model is an (imperfect) reflection of the real world. Over the years, model developments have started with simple descriptions of the 'mean' effects averaged over many individuals in a population. Of course, there is no 'average' animal, as populations consist of heterogeneous individuals. It is only natural that capture-recapture models have become more complex, in order to try and capture some of this heterogeneity—the same driving force as in the closed population models, but with more types of parameters and behaviour that is subject to individual variability. Stratification variables that operate on groups of animals (e.g. sex, age, year) have been introduced to try and account for part of the heterogeneity in responses.

Modern software now allows us to model individual animal covariates. However, there is a danger in taking this too far. For example, if individual covariates are needed to predict future behaviour, what hope is there in managing populations at an aggregate level? Paradoxically, we need to avoid paralysis from too much information (Cooch *et al.*, 2002)!

We note that most of the modelling effort in the CJS framework has followed marked individuals only, and it assumes independence among individual fates. Tools to investigate the whole issue of density dependence and dependence upon the actions of other individuals are not yet readily available. Models that estimate abundance (e.g. Jolly-Seber models) are available, but the feedback loop between abundance and subsequent parameters has not yet been complete. Modelling interactions between individuals is an order more complex; n individuals have about n^2 interactions among pairs of individuals. It is clear that the current approaches for building models will not work.

Another important research question involves trade-offs between competing courses of action. Many capture-recapture studies are pure observational studies and teasing out trade-offs is difficult. Furthermore, the results can always be challenged as mere association rather than causation. Better experimental designs, where demands can be deliberately modified and subsequent behaviour studied, need to be developed.

3.2 Individual covariates

Covariates can apply at two levels. First, covariates that apply at the sampling occasion and are applicable to all animals (e.g. survival probability as a function of winter severity) are relatively easy to fit with current software, and the methodology is well developed. Individual based covariates, however, are more problematical. Such models have been developed for closed populations and CJS studies, and these condition upon the covariates observed in the study. Unconditional inference is more difficult as the distribution of covariates in the unobserved population must also be modelled. Additional work needs to be done with individual covariate models in the JS and robust design models. McDonald & Amstrup (2002) use covariates in the JS models and estimated population size at each sampling point using a Horvitz-Thompson type of estimator, but this approach does not fully integrate recruitment. Kendall (2000) has incorporated individual covariates for the robust design at the secondary (closed) sampling occasions, but not in the primary occasions. The biggest problem for these models is how to model the covariates for the unobserved animals prior to their initial capture.

Individual-based, time varying covariates are the next level of complexity.

Unfortunately there have been no readily identifiable papers in this area, but the theory and methods from the medical literature should provide some guidance.

A biologically interesting covariate model is that of a dependence of survival upon density. If density (or abundance) is known from external sources (e.g. a separate census of the population), then this is easily incorporated into CJS models using the methods above. For example, Catchpole *et al.* (2000) included the known population sizes of Soay sheep in the previous year when modelling the survival over winter. The situation becomes more difficult in JS models where abundance (or density) is an unknown parameter of the model and makes the likelihood function quite complex, as shown by Barker *et al.* (2002). Again further research and model development will be needed here.

3.3 Bayesian methods

The Gibbs sampler and related developments have revolutionized the application of Bayesian methods. Modellers are much less inhibited in their choice of prior distributions and the complexity of the models to be fitted. However, there are several practical issues that need resolving.

First, adequate computational machinery needs to be developed that will make fitting Bayesian methods as 'painless' as current methods. Given the large amount of computation involved in the Gibbs sampler, this may require the next generation of computers. Next, large complex models with many parameters and sparse data may suffer from both non-identifiability and non-estimability. The former is a structural problem with the model that no amount of data will resolve, while the latter may be an artefact of a particular dataset. Catchpole & Morgan (1997) have developed methods to identify problems with the non-Bayesian methods. In the Bayesian methods, these problems 'don't exist'. All parameters are always identifiable and estimable because, in the worst case, the final posterior is identical to the prior distribution if the data provide no information about the parameter. Diagnostic methods need to be developed for Bayesian methods to identify parameters whose posterior distribution is driven largely by the prior. On a related note, diagnostics are also needed to identify which posteriors are sensitive to the choice of the prior distribution. It is not feasible to investigate individually each of several hundred parameters.

The Bayesian approach can also be used for 'model simplification' in much the same way as the random effect models (next section) have been used. This has yet to be explored, but is a very simple extension of Bayesian methods.

3.4 Random effects models

Random effects models introduce model simplification to one problem of parameter proliferation. Replacing 40 yearly survival rates by the mean survival rate and the variation around the rate is an intermediate model between having a large number of survival parameters and a single parameter that is assumed constant over time. This model provides shrinkage estimates of the random parameter by shrinking the MLE towards zero and estimates the effective number of parameters needed in determining the AIC and related indices for model selection.

Not surprisingly, the previous models are close relatives to full Bayesian models, and are computationally more feasible when examining a large number of models. Perhaps a good strategy would be to identify a small set of candidate models using

these methods before attacking the problem in a fully Bayes' context. An obvious area of research is a detailed comparison of the random effects and Bayesian approaches to identify under what conditions they give similar and dissimilar results.

Random effects models can also be used to model heterogeneity among animals rather among parameters. Pledger & Schwarz (2002) develop a non-parametric approach where a simple two-support point distribution effectively accounts for most of the heterogeneity in animal survival. Again, a comparison between this approach and the Bayesian approach is needed.

Some potential problems to be overcome with the random effect models are technical and operational. The big question is why should a practioner use them? What additional information is gleaned from these models? The methods presented can be generalized to other parameters such as population growth. But, unlike survival rates, population growth consists of the net effect of two distinct processes—recruitment and survival. Is it sensible to model this joint process by a single random variable? Perhaps it is more sensible to model simultaneously both processes by a bivariate random variable.

3.5 *Meta-population and multi-stratum models*

Multi-strata models present a unified method to dealing with transition data where the definition of states can be quite general. They were originally developed to deal with movement data, but changes in weight classes, breeding status, and the type of recapture/resighting can also be formulated. The number of parameters can become large very quickly in these models, particularly if a general transition matrix allowing for arbitrary movements among states is allowed. Again, these type of models may profit from 'simplification', i.e. modelling movement rates as functions of covariates such as the distance between areas. There is also a fixed amount of information in the data so that, as mentioned before, diluting the information over many parameters leads to poor inference on any one parameter. Simplification can also take place across time. For example, can Burnham's (2001) random effects models be generalized to allow yearly movement rates to be random fluctuations around an overall mean? This will require a multivariate approach as the movement rates are obviously constrained to sum to one within each year.

The session also identified some technical issues that need to be resolved in the next few years. The likelihood surfaces are very flat, and it is relatively easy to get stuck at a local maximum rather than the global maximum. The flatness of the likelihood is also related to identifiability and estimability concerns so that better diagnostics need to be developed. Many of the parameters may have estimates at the boundary of the parameter space. How do these affect the model selection criterion? Once a model is developed, how is the goodness of fit assessed? At the moment, there are no general measures, as have been developed for the CJS models.

Several methodological issues also arose. First, a generalization of the Arnason (1972, 1973) models that incorporate abundance estimation need to be available in software. Second, the Stratified-Petersen estimator (Plante *et al.*, 1988; Schwarz & Taylor, 1989) is not often used in CJS studies, but it has applications for bird studies when abundance is of interest. Lastly, is there a way to generalize the Pradel (1996) methods of reading the capture histories backwards in a similar fashion for movement studies?

3.6 Formal experiments

The design and analysis of experiments is a well developed field in many experimental sciences. However, many bird studies are purely observational in nature and few are manipulative. This leads to the obvious mantra 'No causation with manipulation'. Manipulative experiments are often difficult to perform and so the gold-standard many not be an attainable goal. However, medical studies are faced with similar problems, and this field does much more with meta-analysis and comparing results across many different studies. These techniques could prove to be an important intermediate step before formal experiments.

One of the key problems encountered in the analysis of formal experiments is the distinction between the observational unit (the bird) and the experimental unit. In some cases these do not coincide, e.g. if the experimental unit is a colony. This gives rise to several different 'error' terms in the formal ANOVA of experiments, and these do not seem to have any counterpart in many bird studies. Generalizing the CJS model to include multiple sizes of experimental units will be difficult—the field of generalized linear mixed models is large and complex—and we suspect that the Bayesian methods will be the general way to analyse these types of experiments.

Pseudo-replication (Hurlbert, 1984) is also an important issue. For example, studies where sets of birds are taken from one urban and one rural area to compare their survival rates are examples of pseudo-replication. Here, inference is limited to comparison between these two specific population and not to urban/rural areas in general. Again, the subject of meta analyses, where many studies of limited size are combined, needs more attention.

3.7 Population growth

Rather than estimating abundance directly, which is a difficult problem because of biases introduced by heterogeneity in catchability, it may be easier to measure population change, i.e. growth or decline. One traditional tool has been Leslie matrices where the dominant eigenvalue indicates the asymptotic properties of the population. However, many populations are far from asymptopic, and the time-specific population growth rates can be estimated using the CJS models.

As was shown in the conference many times, these methods are equivalent to the ordinary Jolly-Seber model, and so the same assumptions must be made. In particular, changes in study area definition are completely confounded with changes in population size. There is a real danger in uncritically using CJS studies, where birds can be marked and released in a haphazard fashion, for population growth studies where modelling the latter process is what determines population growth.

The current methods allow fairly straightforward estimation of population growth rates, but how these depend upon other characteristics of the population is also of interest. Parameters such as elasticity and density dependence need to be incorporated into the analysis methods.

Can the Leslie matrix approach and the CJS approaches be unified? The missing link in capture-recapture studies is that the new entrants to the population are not identified as to their origin, e.g. by age of the parents. If this information could be captured, then it should be possible to construct a version of the CJS model that also estimates fecundity (Nichols & Hines, 2002). It is also possible to combine mark-recapture with census information (Besbeas *et al.*, 2002) to estimate all the necessary parameters.

4 Challenges for the future

Several challenges are apparent. The technical challenges have been presented in the above sections and will not be repeated here. However, there are several cross-sessional and cross-model challenges to be addressed.

The EURING conferences are the premiere conferences for new developments in capture-recapture methodology. Indeed, much of recent developments in these methods has been driven by past conferences. However, capture-recapture methods are used by other disciplines as well, and the challenge is to transfer the technology in both directions. For example, studies in mammals often use capture-recapture methods, but how do these scientists learn of the new advances? In the other direction, a vast literature has developed in using multi-list methods in medical problems (e.g. disease ascertainment), but few of these developments have appeared in the EURING conferences.

Where does a scientist begin in designing a study? What are the trade-offs between using simple banding methods and more complex radio-telemetry studies. For example, approximately 800 radio-tags with 100% detection probabilities give approximately the same precision for survival estimates as 20 000 simple tags with 20% detection probabilities! How should effort be allocated among recaptures and resightings when both types of recoveries are available? How much total effort is needed? Should studies be done that have only a 30% chance of detecting an important effect or that yield estimates with coefficients of variation approaching 100%? Despite the sophistication in tools for analysing completed studies, the planning process has been neglected. A comprehensive guide to the trade-offs involved in designing these studies needs to be developed.

Density effects are conspicuously absent from the modelling efforts. How can density be incorporated as a covariate or parameter into the models? Models with internal feedback loops are notoriously difficult to model. Much of the fisheries literature is concerned with stock assessments, which include a stock-recruitment component, and the lessons from these studies should be incorporated into bird studies.

Capture-recapture studies are used for large and small scale studies. However, the objectives of these types of studies are quite different and it would be unusual if the same study design and trade-offs are applicable to both types of designs. Some specialized designs need to be developed for the different scales of studies.

In conclusion, we see that the explosion of papers on estimating animal population parameters that have appeared in the last 20 years reflects the importance of the subject, the increased computing power available, and the increased statistical sophistication of the practitioners. As noted in Schwarz & Seber (1999), many statistical methods have been used to attack these problems so that the training of future scientists in this subject area should not concentrate only on population methodology. The training needs to be broadly based regarding statistics and computing.

REFERENCES

ALHO, J. M. (1990) Logistic regression in capture-recapture models, *Biometrics*, 46, pp. 623-635.

ARNASON, A. N. (1972) Parameter estimation from mark-recapture experiments on two populations subject to migration and death, *Researches in Population Ecology*, 13, pp. 97-113.

ARNASON, A. N. (1973) The estimation of population size, migration rates, and survival in a stratified population, *Researches in Population Ecology*, 15, pp. 1-8.

BARKER, R. J. (1999) Joint analysis of mark-recapture, resighting and ring-recovery data with age-dependence and marking-effect, *Bird Study*, 46, Supplement, pp. 82-91.

BARKER, R., FLETCHER, D. & SCOFIELD, P. (2002) Measuring density dependence in survival from a mark-recapture study. *Journal of Applied Statistics*, this issue.

BECKER, N. G. (1984) Estimating population size from capture-recapture experiments in continuous time, *Australian Journal of Statistics*, 26, pp. 1-7.

BESBEAS, T., FREEMAN, S., MORGAN, B. J. T. & CATCHPOLE E. A. (2002), Integrated ring-recovery and census data analysis. *Biometrics*, in preparation.

BROOKS, S. P., CATCHPOLE, E. A. & MORGAN, B. J. T. (2000a) Bayesian animal survival estimation, *Statistical Science*, 15, pp. 357-376.

BROOKS, S. P., CATCHPOLE, E. A., MORGAN, B. J. T. & BARRY, S. C. (2000b) On the Bayesian analysis of ring-recovery data, *Biometrics*, 56, pp. 951-956.

BROOKS, S. P., CATCHPOLE, E. A. & MORGAN, B. J. T. (2001) Bayesian Methodology for Analysing Ringing Data. *Journal of Applied Statistics*, this issue.

BROWNIE, C., HINES, J. E., NICHOLS, J. D., POLLOCK, K. H. & HESTBECK, J. B. (1993) Capture-recapture studies for multiple strata including non-Markovian transition probabilities, *Biometrics*, 49, pp. 1173-1187.

BUCKLAND, S. T., BURNHAM, K. P. & AUGUSTIN, N. H. (1997) Model selection: an integral part of inference, *Biometrics*, 53, pp. 603-618.

BURNHAM, K. P. (2001) Random effects models in ringing and capture-recapture studies, *Journal of Agricultural, Biological, and Environmental Statistics*, to appear.

BURNHAM, K. P. & ANDERSON, D. R. (1998) *Model Selection and Inference: a Practical Information-theoretic Approach* (New York, Springer-Verlag).

BURNHAM, K. P. & WHITE, G. C. (2002) Evaluation of some random effects methodology applicable to bird ringing data, *Journal of Applied Statistics*, this issue.

BURNHAM, K. P., ANDERSON, D. R. & WHITE, G. C. (1995) Selection among open population capture-recapture models when capture probabilities are heterogeneous, *Journal of Applied Statistics*, 22, pp. 611-624.

CATCHPOLE, E. A. & MORGAN, B. J. T. (1996) Model selection in ring-recovery models using score tests, *Biometrics*, 52, pp. 664-672.

CATCHPOLE E. A. & MORGAN, B. J. T. (1997) Detecting parameter redundancy, *Biometrika*, 84, pp. 197-196.

CATCHPOLE, E. A., MORGAN, B. J. T., COULSON, T. N., FREEMAN, S. N. & ALBON, S. D. (2000) Factors influencing Soay sheep survival. *Applied Statistics*, 49, 453-478.

CHAO, A., LEE, S.-M. & JENG, S.-L. (1992) Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal, *Biometrics*, 48, pp. 201-216.

COOCH, E. G., CAM, E. & LINK, W. (2002) Occam's Revenge: levels of analysis in evolutionary ecology—where to next? *Journal of Applied Statistics*, this issue.

CORMACK, R. M. (1964) Estimates of survival from the sighting of marked animals, *Biometrics*, 51, pp. 429-438.

CORMACK, R. M. (1989) Log-linear models for capture-recapture, *Biometrics*, 45, pp. 395-413.

DUPUIS, J. A. (1995) Bayesian estimation of movement and survival probabilities from capture-recapture data, *Biometrika*, 82, pp. 761-772.

EVANS, M. A., BONETT, D. G. & MCDONALD, L L. (1994) A general theory for modeling capture-recapture data from a closed population, *Biometrics*, 50, pp. 396-405.

FREEMAN, S. N. (1990) Statistical analysis of avian breeding and survival. PhD Thesis, University of Kent, Canterbury, Kent, UK.

FREEMAN, S. N. & MORGAN, B. J. T. (1992) A modeling strategy for recovery data from birds ringed as nestlings, *Biometrics*, 48, pp. 217-235.

GARTHWAITE, P. H., YU, K. & HOPE, P. B. (1995) Bayesian analysis of a multiplere-capture model, *Communications in Statistics—Theory and Methods*, 24, pp. 2229-2247.

GOULD, W. R. & POLLOCK, K. H. (1997) Catch-effort estimation of population parameters under the robust design, *Biometrics*, 53, pp. 207-216.

HOETING, J. A., MADIGAN, D., RAFTERY, A. E. & VOLINSKY, C. T. (1999) Bayesian model averaging: a tutorial, *Statistical Science*, 14, pp. 382-417.

HUGGINS, R. M. (1989) On the statistical analysis of capture experiments, *Biometrika*, 76, pp. 133-140.

HURLBERT, S. H. (1984) Pseudoreplication and the design of ecological field experiments, *Ecological Monographs*, 54, pp. 187-211.

HWANG, W.-D. & CHAO, A. (1995) Quantifying the effects of unequal catchabilities on Jolly-Seber estimators via sample coverage, *Biometrics*, 51, pp. 128-141.

JOLLY, G. M. (1965) Explicit estimates from capture-recapture data with both death and immigration-stochastic model, *Biometrika*, 52, pp. 225-247.

KENDALL, W. L (2000) The robust design for capture-recapture studies: analysis using Program MARK, Hungarian Symposium.

KENDALL, W. L. & NICHOLS, J. D. (1995) On the use of secondary capture-recapture samples to estimate temporary emigration and breeding proportions, *Journal of Applied Statistics*, 22, pp. 751-762.

LEBRETON, J.-D. & PRADEL, R. (2002) Multistratum recapture models: modeling incomplete individual histories, *Journal of Applied Statistics*, this issue.

LEBRETON, J.-D., ALMERAS, T. & PRADEL, R. (1999) Competing events, mixture of information, and multi-strata recapture models, *Bird Study*, 46, pp. S39-S46.

LEBRETON, J.-D., BURNHAM, K. P., CLOBERT, J. & ANDERSON, D. R. (1992) Modeling survival and testing biological hypotheses using marked animals. A unified approach with case studies, *Ecological Monographs*, 62, pp. 67-118.

LEE, A. J. & SEBER, G. A. F. (2001) Residuals, *Encyclopedia of Environmetrics*, to appear.

LINDBERG, M. S., KENDALL, W. L., HINES, J. E. & ANDERSON, M. G. (2001) Combining band recovery data and Pollock's robust design to model temporary and permanent emigration, *Biometrics*, 57, pp. 273-281.

LLOYD, C. J. & YIP, P. (1991) A unification of inference for capture-recapture studies through martingale estimating functions. *In:* V. P. GODAMBE (Ed), *Estimating Equations*, pp. 65-68 (Oxford, Clarendon Press).

MCDONALD, T. L., & AMSTRUP, S. C. (2001) Estimation of population size using open capture-recapture models, *Journal of Agricultural, Biological, and Environmental Statistics*, 6, pp. 206-220.

NICHOLS, J. D. & HINES, J. E. (2002) Approaches for the direct estimation of lambda and related parameters using capture-recapture data, *Journal of Applied Statistics*, this issue.

NICHOLS, J. D. & POLLOCK, K. H. (1990) Estimation of recruitment from immigration versus in situ reproduction using Pollock's Robust design, *Ecology*, 71, pp. 21-26.

NICHOLS, J. D., SAUER, J. R., POLLOCK, K. H. & HESTBECK, J. B. (1992) Estimating transition probabilities for stage-based population matrices using capture-recapture data, *Ecology*, 73, pp. 306-312.

NORRIS, J. L. & POLLOCK, K. H. (1996) Nonparametric MLE under two closed capture-recapture models with heterogeneity, *Biometrics*, 52, pp. 639-649.

NORTH, P. M. & MORGAN, B. J. T. (1979) Modelling heron survival using weather data, *Biometrics*, 35, pp. 667-681.

OTIS, D. L., BURNHAM, K. P., WHITE, G. C. & ANDERSON, D. R. (1978) Statistical inference from capture data on closed animal populations, *Wildlife Monographs*, 62, pp. 1-135.

PLANTE, N., RIVEST, L. P. & TREMBLAY, G. (1998) Stratified capture-recapture of the size of a closed population, *Biometrics*, 54, pp. 47-60.

PLEDGER, S. (2000) Unified maximum likelihood estimates for closed capture-recapture models using mixtures, *Biometrics*, 56, pp. 434-442.

PLEDGER, S. & EFFORD, M. (1998) Correction of bias to heterogeneous capture probability in capture-recapture studies of open populations, *Biometrics*, 54, pp. 888-898.

PLEDGER, S. & SCHWARZ, C. J. (2002) Modelling heterogeneity of survival as a random effect using finite mixtures, *Journal of Applied Statistics*, this issue.

POLLOCK, K. P. (2002) Modeling and inference using individual covariates, *Journal of Applied Statistics*, this issue.

POLLOCK, K. H. (1982) A capture-recapture design robust to unequal probabilities of capture, *Journal of Wildlife Management*, 46, pp. 757-760.

POLLOCK, K. H., HINES, J. E. & NICHOLS, J. D. (1984) The use of auxiliary variables in capture-recapture and removal experiments, *Biometrics*, 40, pp. 329-340.

PRADEL, R. (1996) Utilization of capture-mark-recapture for the study of recruitment and population growth rates, *Biometrics*, 52, pp. 371-377.

SCHWARZ, C. J. (2002) Real and quasi-experiments in bird ringing studies: lessons to be learned from the past and sister disciplines, *Journal of Applied Statistics*, this issue.

SCHWARZ, C. J. & DEMPSON, J. B. (1994) Mark-recapture estimation of a salmon smolt population, *Biometrics*, 50, pp. 98-108.

SCHWARZ, C. J. & SEBER, G. A. F. (1999) Estimating animal abundance: review III, *Statistical Science*, 14, pp. 427-456.

SCHWARZ, C. J. & STOBO, W. T. (1997) Estimating temporary migration using the robust design, *Biometrics*, 53, pp. 178-194.

SCHWARZ, C. J. & TAYLOR, C. G. (1998) The use of the Stratified-Petersen estimator in fisheries management: estimating the number of pink salmon (*Oncorhynchus Gorbuscha*) spawners in the Fraser River, *Canadian Journal of Fisheries and Aquatic Sciences*, 55, pp. 281-297.

SCHWARZ, C. J., SCHWEIGERT, J. & ARNASON, A. N. (1993) Estimating migration rates using tag recovery data, *Biometrics*, 49, pp. 177-194.

SEBER, G. A. F. (1965) A note on the multiple recapture census, *Biometrika*, 52, pp. 249-259.

SEBER, G. A. F. (1973) *The Estimation of Animal Abundance and Related Parameters*. (London, Edward Arnold).

SEBER, G. A. F. (1982) *The Estimation of Animal Abundance and Related Parameters*, 2nd edn (London, Edward Arnold).

SEBER, G. A. F. (1986) A review of estimating animal abundance, *Biometrics*, 42, pp. 267-292.

SEBER, G. A. F. (1992) A review of estimating animal abundance. II, *International Statistical Review*, 60, pp. 129-166.

UNDERHILL, L. G. (1990) Bayesian estimation of the size of a closed population, *The Ring*, 13, pp. 235-254.

VOUNATSOU, P. & SMITH, A. F. M. (1995) Bayesian analysis of ring-recovery data via Markov chain Monte Carlo simulation, *Biometrics*, 51, pp. 687-708.

YIP, P. (1989) An inference procedure for a capture and recapture experiment with time-dependent capture probabilities, *Biometrics*, 45, pp. 471-479.