New developments in change-in-ratio and index-removal methods, with application to snow crab (*Chionoecetes opilio*)

Chiu-Lan Chen, John M. Hoenig, Earl G. Dawe, Cavell Brownie, and Kenneth H. Pollock

Abstract: Change-in-ratio and index-removal estimates of population size and related parameters can be obtained when a survey is conducted before and after a fishery and the magnitude and composition of the landings are determined. The former looks at how a known, selective removal affects the composition of the population. The latter looks at how a known removal affects catch rate. If spatial pattern remains relatively stable between the two surveys, increased precision can be obtained by occupying the same sampling stations in the two surveys rather than rerandomizing locations for the second survey. We derive expressions for the biases and variances of the estimates for the case where stations are reoccupied. For snow crab (*Chionoecetes opilio*) in Newfoundland, reuse of stations led to reductions in standard errors of 33–47% from the usual rerandomized-station design. An assumption of the index-removal method is that all animals have the same probability of capture in the surveys. This is not true when survey sampling gear is size selective. The bias can be minimized by making separate estimates for each size-class. However, when large animals have greater catchability than smaller animals, we can achieve greater precision by utilizing information on relative catchability. We propose that separate estimates for each size-class first be computed. If the estimated catchability coefficients show an increasing trend with size, these may then be constrained to increase monotonically with size. It could also be assumed that the catchability coefficients must be a specified function of size. We conclude that these methods may have considerable value for assessing invertebrate stocks.

Résumé : On peut obtenir des estimations dla tailationres congemenprélèvur lorfectue un relevé avant et après une pêche, et que l'on connaît l'importance et la composition des prises. Dans le premier cas, on examine de quelle manière un prélèvement sélectif connu influe sur la composition de la population. Dans le second, on examine de quelle manière un prélèvement connu influe sur le taux de capture. Si la distribution spatiale reste relativement stable entre les deux relevés, on peut avoir une meilleure précision en occupant les deux fois les mêmes stations d'échantillonnage plutôt qu'en choisissant de nouveau les sites de manière aléatoire pour le second relevé. Nous tirons des expressions des biais et variances des estimations pour les cas où l'on a occupé les mêmes stations les deux fois. Pour les crabes des neiges (*Chionoecetes opilio*) de Terre-Neuve, l'utilisation des mêmes stations a permis de réduire l'erreur-type de 33 à 47 % par comparaison avec les techniques habituelles de randomisation du choix des stations. Une des hypothèses de la méthode du prélèvement indicateur est que tous les individus ont la même probabilité d'être capturés au cours des relevés. Ce n'est pas le cas lorsque les engins d'échantillonnage sont sélectifs sur le plan de la taille. On peut réduire le biais à un minimum en effectuant des estimations séparées pour chaque classe de taille. Cependant, quand les gros spécimens ont une plus grande capturabilité que les petits, on peut arriver à une meilleure précision en utilisant l'information sur la capturabilité relative. Nous proposons de calculer en premier des estimations séparées pour chaque classe de taille. Si les coefficients de capturabilité estimatifs montrent que la tendance croît avec la taille, on peut les contraindre pour obtenir une augmentation monotone avec la taille. On peut aussi postuler que les coefficients de capturabilité sont une fonction donnée de la taille. Nous en concluons que ces méthodes peuvent être précieuses pour évaluer les stocks d'invertébrés. [Traduit par la Rédaction]

Introduction

In this paper, we consider how population size, exploitation rate, and related parameters can be estimated from survey data collected just before and just after a known (or estimated) harvest. It is assumed that the time between the start of the first and the end of the second survey is sufficiently short that the population can be considered closed except for the removals. There are three types of estimators that can be used with these

C.-L. Chen. Biomathematics Program, North Carolina State University, Raleigh, NC 27695, U.S.A.

J.M. Hoenig[1] and Earl G. Dawe. Department of Fisheries and Oceans, P.O. Box 5667, St. John's, NF A1C 5X1, Canada.

C. Brownie and K.H. Pollock. Department of Statistics, North Carolina State University, Raleigh, NC 27695, U.S.A.

Correct citation: Chen, C.-L., Hoenig, J.M., Dawe, E.G., Brownie, C., and Pollock, K.H. 1998. New developments in change-in-ratio and index-removal methods, with application to snow crab (*Chionoecetes opilio*). *In* Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. *Edited by* G.S. Jamieson and A. Campbell. Can. Spec. Publ. Fish. Aquat. Sci. 125. pp. 49–61.

[1] Author to whom all correspondence should be sent at the following address: Virginia Institute of Marine Science, P.O. Box 1346, Gloucester Point, VA 23062, U.S.A.

data: change-in-ratio, index-removal, and removal methods. The methods are simple and have been known for at least 40 years but, surprisingly, they are not well known to fishery scientists and much work remains to be done to elucidate their statistical properties.

We begin by reviewing the available methods. We then show that increased statistical efficiency can generally be achieved by using the same sampling locations for both surveys rather than rerandomizing the stations for the second survey. The improvements in precision are studied by simulation and in an example involving snow crab (*Chionoecetes opilio*) from St. Mary's Bay, Newfoundland, Canada. We also show that, when information exists on the relative catchability of different groups in the population (e.g., size groups), this information can be incorporated in the index-removal estimation scheme.

Review of models

Change-in-ratio estimation

Change-in-ratio methods make use of the change in population composition caused by a known, selective removal. For example, if a known removal of males from a population causes the proportion that is male to change a great deal, then the population must be relatively small whereas, if the removal only changes the sex ratio by a small amount, then the population must be relatively large. For the change-in-ratio method, the population must be divided into two classes (say, x- and y-type) on the basis of size, sex, or another distinguishing feature and the removal must be selective for one of the classes. Formally, we have

$$[1] \quad \hat{N}_{\text{cir}} = \frac{\hat{R}_x - \hat{P}_2 \hat{R}}{\hat{P}_1 - \hat{P}_2}$$

$$[2] \quad \hat{X}_{\text{cir}} = \hat{P}_1 \hat{N}_{\text{cir}}$$

where $\hat{N}_{\text{cir}}$ is the change-in-ratio estimate of initial population size ($x + y$ types combined), $\hat{X}_{\text{cir}}$ is the estimate of the initial number of x-type animals, $\hat{R}$ is the estimated number of animals removed, $\hat{R}_x$ is the estimated number of x-type animals (e.g., males) removed, $\hat{P}_1$ is the estimated proportion that is x-type during the survey prior to the selective removal, and $\hat{P}_2$ is the estimated proportion that is x-type in the post-removal survey. Equations 1 and 2 are the maximum likelihood estimators when the estimates of the proportions and removals are maximum likelihood estimates (Dawe et al. 1993). Change-in-ratio estimation has been known at least since Kelker (1940).

It is possible to generalize change-in-ratio estimators to allow for more than two classes in the population and for additional surveys and removals. It is also possible to relax the assumptions. Recent developments are presented by Udevitz and Pollock (1991, 1995) and reviewed by Pollock and Hoenig (1998).

In order to obtain (asymptotically) unbiased estimates of the sizes of the x- and y-type populations, the probability of capture in a survey must be equal for the two classes. However, if only one class is harvested then the estimate of population size for that class is unbiased even if the proportions estimated in the surveys are biased due to differential catchability of the two types of animals (see Seber 1982 and Pollock et al. 1985 for discussions).

Index-removal estimators

The index-removal method makes use of the decline in relative abundance due to a known removal. If catch rate (catch per unit of sampling effort) is proportional to animal abundance, and if a known removal causes the catch rate to decline by a specified proportion P, then the removal is equal to $100\,P$ % of the population. For example, if the catch rate is 10 before the removal of 300 animals and is 7 after the removal is made, then we calculate that the removal of 300 animals resulted in a loss of $(10 - 7)/10 = 3/10$ of the population. Thus, the population size (before the removal) must have been 1000 animals. More formally, if $E(c_{x1})$ and $E(c_{x2})$ are the expected values of the observed catch rates of a component, x, of the population before and after the removal, respectively, and if R_x is the number of x-type animals removed, then the initial size of the population of x-type animals is given by

$$X_{\text{ir}} = \frac{E(c_{x1})\, R_x}{E(c_{x1}) - E(c_{x2})}$$

and an estimate can be obtained as

$$[3] \quad \hat{X}_{\text{ir}} = \frac{\hat{c}_{x1}\, \hat{R}_x}{\hat{c}_{x1} - \hat{c}_{x2}}$$

where $\hat{c}_{x1}$ and $\hat{c}_{x2}$ are the estimated mean catch rates in the first and second surveys, respectively, and $\hat{R}_x$ is the removal or an estimate of the removal.

This approach is well known in the wildlife literature (Petrides 1949; Eberhardt 1982; Seber 1982; Roseberry and Woolfe 1991) but has received little attention in the fisheries literature (Dawe et al. 1993). Seber (1982) and Routledge (1989) discuss the statistical theory in detail. In particular, Routledge (1989) generalized the approach to include J removal periods and $J + 1$ surveys. As developed above, eq. 3 is a moment estimator. Eberhardt (1982) showed it is also the maximum likelihood estimator when the catches in the two surveys are independent Poisson random variables. Dawe et al. (1993) pointed out that eq. 3 is the maximum likelihood estimator for population size whenever the estimates of the catch rates and removal are maximum likelihood. See also Hoenig and Pollock (1998) for a review.

For the simplest case described above, the assumptions of the method are as follows: (*i*) the population is closed except for the removals which are known exactly, and (*ii*) all animals have the same probability of capture which does not change from survey to survey. It is easily verified that heterogeneity of capture probabilities can introduce bias. Suppose, for example, that the population is composed of 500 males and 500 females, that males have a catchability coefficient of 0.01 whereas females have a catchability coefficient of 0.005 (i.e., half that of the males), and that 300 males and 100 females are removed from the population between the time of the two surveys. In the first survey we would expect to catch $0.01 \times 500 = 5$ males if one randomly placed unit of sampling effort is expended. In the second survey we would expect to catch $0.01 \times (500 - 300) = 2$ males. Thus, the calculated size of the initial population of males would be

$$X_{\text{males}} = \frac{5}{5-2} \times 300 = 500$$

which is what we want. Similarly, the size of the female population would be calculated to be 500, as desired. However, suppose it was not known that males have a different catchability coefficient than females and the size of the total population was calculated from combined data on males and females. In the first survey, one would expect that 7.5 animals (5 males + 2.5 females) would be caught with one randomly placed unit of sampling effort. In the second survey, the expectation would be 2 + 2 = 4 animals. Consequently, the calculated population size would be

$$N = X_{\text{males}} + X_{\text{females}} = \frac{7.5}{7.5-4} \times 400 = 857$$

instead of the actual value of 1000. Note that the heterogeneity of capture probabilities is a problem because the removal was selective with respect to capture probabilities (i.e., proportionately more of the males were removed than of the females). The problem of heterogeneity can be avoided by making separate estimates for each component of the population provided that capture probabilities within a population component are homogeneous (and do not change over time) and provided that the removals are known by component.

Removal estimators

The simplest removal estimator is obtained by dividing the fishing season in two parts with the fishing effort in the first part equal to that in the second part. The population is closed except for the removals and catch rate is assumed proportional to abundance. Thus, we would expect the catch in the second part of the season to be smaller than that in the first part because abundance was reduced in the first part by harvest. The maximum likelihood estimator is

$$\hat{N}_r = \frac{\hat{R}_1 \hat{R}_1}{(\hat{R}_1 - \hat{R}_2)} = \frac{\hat{R}_1{}^2}{(\hat{R}_1 - \hat{R}_2)}$$

where $\hat{R}_1$ and $\hat{R}_2$ are the catches (or, more generally, estimated catches) in the first and second parts of the season, respectively. This estimator is very similar in structure to the index-removal estimator. However, in the index-removal estimator the relative abundance is measured by the pre- and post-season surveys and the removal is done separately by the fishery. In contrast, in the removal estimator, the fishery information is used both to measure relative abundance and to deplete the population. Pre- and post-season surveys don't have to be conducted to use the removal method but catch and effort data for each subdivision of the fishing season must be obtained.

The removal method dates back to the work of Zippin (1956), Leslie and Davis (1939), and DeLury (1947). Modern generalizations are reviewed by Seber (1982).

Combining methods

It is possible to combine methods. Routledge (1989) combined the removal and index-removal methods using a likelihood framework. Dawe et al. (1993) combined the change-in-ratio and index-removal methods by computing a weighted mean. Recently, Chen (1995) combined change-in-ratio, index-removal, and removal estimators in a single likelihood function.

When more than one method is appropriate, increased efficiency can be achieved by combining methods. However, use of an estimator when the assumptions are not met will generally lead to biased results. Assessment of the assumptions is thus important. The index-removal method requires a stronger assumption (constant gear efficiency over time) than the change-in-ratio estimator (equal gear efficiencies not necessarily constant over time or constant ratio of gear efficiencies over time if only one group is exploited).

Estimation of exploitation rate, catchability coefficient, and fishing power of the gear

In addition to estimating population size (for various subsets of the population), it is also possible to estimate exploitation rate, catchability coefficient, and fishing power of the sampling gear. The exploitation rate, u, is the fraction of the population present at the beginning of the study that is harvested. It can be estimated by dividing the harvest or estimated harvest by the estimated initial population, $\hat{X}_{\text{cir}}$ or $\hat{X}_{\text{ir}}$. Thus, the estimated exploitation rate by the index-removal method is (Eberhardt 1982)

$$\hat{u}_{\text{xir}} = \frac{\hat{R}_x}{\hat{X}_{\text{ir}}} = \frac{\hat{c}_{x1} - \hat{c}_{x2}}{\hat{c}_{x1}}$$

which does not require that the total removal be known. The estimated exploitation rate by the change-in-ratio method is

$$\hat{u}_{\text{xcir}} = \frac{\hat{R}_x}{\hat{X}_{\text{cir}}} = \frac{\hat{f}(\hat{P}_1 - \hat{P}_2)}{\hat{f}\hat{P}_1 - \hat{P}_1 \hat{P}_2}$$

for which it is necessary to know the fraction of the harvest which is x-type (f) but not the magnitude of the harvest.

The catchability coefficient, q, is the fraction of the population taken by one randomly placed unit of sampling effort when the fraction taken is small (e.g., less than 2%; see Ricker 1975). It can be estimated by dividing the initial catch rate by the estimated initial population size. The catchability coefficient can be expressed as

$$[4] \quad q = \frac{\rho \phi a}{\phi A} = \frac{\rho a}{A}$$

where ϕ is the average density of animals (per unit area), A is the area inhabited by the stock, a is the area fished by the sampling gear (per unit of sampling effort), and ρ is the proportion of the animals encountering the sampling gear that is retained. The denominator is simply the total population size while the numerator is the expected catch from one unit of sampling effort. The catchability coefficient is population-specific because its definition includes the area inhabited by the population. If this area is known and if the estimate of q is multiplied by A, then we obtain an estimate of the fishing power of the gear, ρa, which is a property of the sampling gear. Consequently, we can obtain an estimate of the catchability coefficient q^* for a new population inhabiting an area A^* by dividing the estimated fishing power by A^*.

Efficient sampling design

An important feature of the work of Eberhardt (1982) and Dawe et al. (1993) was the recognition that, in practice, catch rates are usually estimated with replicate samples so that the variance can be estimated empirically rather than by resorting to theoretical variances under the assumption of a Poisson, binomial, or hypergeometric distribution of catches as described by Seber (1982). Also, the population proportions are usually estimated by cluster sampling rather than by simple random sampling (Dawe et al. 1993).

Thus, if a standard method of sampling is employed at each of n randomly selected locations at time j (j = 1, 2), then the proportion of animals which is type "x" is estimated by

$$[5] \quad \hat{P}_j = \frac{\sum_{i=1}^{n} x_{ij}}{\sum_{i=1}^{n} z_{ij}}$$

where x_{ij} and z_{ij} are the number of x-type animals and the total number of animals caught, respectively, at the i^{th} location (i = 1, ... n). The variance of this ratio estimator can be estimated by (Cochran 1977)

$$[6] \quad \hat{V}(\hat{P}_j) = \frac{\sum_{i=1}^{n} (x_{ij} - \hat{P}_j z_{ij})^2}{n(n-1)\,\bar{z}_j^2}$$

where $\bar{z}_j$ is the mean of the n observations of z_{ij}.

If the same stations are occupied in the pre- and post-removal surveys, the estimated proportions will not be independent. The covariance of $\hat{P}_1$ and $\hat{P}_2$ can be estimated by

$$[7] \quad \widehat{\mathrm{Cov}}(\hat{P}_1, \hat{P}_2) = \frac{\sum_{i=1}^{n} (x_{i1} - \hat{P}_1 z_{i1})(x_{i2} - \hat{P}_2 z_{i2})}{n(n-1)\,\bar{z}_1 \bar{z}_2} .$$

If the spatial pattern is persistent (areas with a high proportion of x-type animals before the removal tend to have a high proportion after the removal and areas with a low proportion before tend to have a low proportion after the removal) then we would expect the estimated covariance to be positive.

The catch rate for x-type animals in survey j can be estimated by

$$[8] \quad \hat{c}_{xj} = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$$

with variance estimated by the usual formula

$$[9] \quad \hat{V}(\hat{c}_{xj}) = \frac{\sum_{i=1}^{n} (x_{ij} - \hat{c}_{xj})^2}{n(n-1)} .$$

The covariance of the estimated catch rates in the two surveys can be estimated by

$$[10] \quad \widehat{\mathrm{Cov}}(\hat{c}_{x1}, \hat{c}_{x2}) = \frac{\sum_{i=1}^{n} (x_{i1} - \hat{c}_{x1})(x_{i2} - \hat{c}_{x2})}{n(n-1)} .$$

Again, we would expect the estimated covariance to be positive if the spatial pattern is persistent over time (high abundance areas tend to remain high and low abundance areas tend to remain low).

We may also need estimates of $\mathrm{Cov}(\hat{P}_j, \hat{c}_{xk})$ for j = 1, 2 and k = 1, 2. These can be estimated by

$$[11] \quad \widehat{\mathrm{Cov}}(\hat{P}_j, \hat{c}_{xk}) = \frac{\sum_{i=1}^{n} (x_{ij} - \hat{P}_j z_{ij})(x_{ik} - \hat{c}_{xk})}{n(n-1)\,\bar{z}_j} .$$

The total removal from the population, by type of animal, is assumed to be known or estimated. It is convenient to parameterize the removals in terms of the total removals, R, and the fraction, f, that is x-type, such that the removal of x-type animals, R_x, is $R_x = fR$. The sampling design and the resulting variance of the estimated removals will depend on the particular circumstances and thus are not considered here. We assume that the variances have been estimated and that the estimates are independent of the estimates of the proportions and catch rates in the surveys.

All of the estimators are functions of the random variables $\{\hat{P}_1, \hat{P}_2, \hat{c}_{x1}, \hat{c}_{x2}, \hat{f}, \hat{R}\}$ and we can estimate the variances and covariances of these random variables as discussed above. Estimates of the variances of the estimators of population parameters can be obtained using the Taylor's series or delta method (see Seber 1982, pp. 7–8). In general terms, the variance of a function, g, of a set of random variables $\boldsymbol{Y} = \{Y_1, Y_2, ..., Y_p\}$ can be estimated by

$$[12] \quad \hat{V}(g(\boldsymbol{Y})) = \sum_{r=1}^{p} \sum_{s=1}^{p} \frac{\partial g}{\partial Y_r} \frac{\partial g}{\partial Y_s} \widehat{\mathrm{Cov}}(Y_r, Y_s)$$

where the derivatives are evaluated at the parameter estimates.

The variances of the estimates of the initial number of x-type animals are thus found to be

$$[13] \quad \hat{V}(\hat{X}_{\mathrm{ir}}) = \frac{\hat{X}_{\mathrm{ir}}^2 \hat{V}(\hat{c}_{x2}) + \hat{X}_2^2 \hat{V}(\hat{c}_{x1})}{(\hat{c}_{x1} - \hat{c}_{x2})^2} + \frac{\hat{c}_{x1}^2 \hat{V}(\hat{R}_x)}{(\hat{c}_{x1} - \hat{c}_{x2})^2} - \frac{2\hat{X}_{ir}\hat{X}_2 \widehat{\mathrm{Cov}}(\hat{c}_{x1}, \hat{c}_{x2})}{(\hat{c}_{x1} - \hat{c}_{x2})^2}$$

where $\hat{X}_2 = \hat{X}_{\mathrm{ir}} - \hat{R}_x$, and

$$[14] \quad \hat{V}(\hat{X}_{cir}) = \frac{\hat{N}_1^2 \hat{P}_2^2 \hat{V}(\hat{P}_1) + \hat{N}_2^2 \hat{P}_1^2 \hat{V}(\hat{P}_2)}{(\hat{P}_1 - \hat{P}_2)^2} + \frac{\hat{P}_1^2 \hat{R}^2 \hat{V}(\hat{f}) + \hat{P}_1^2 (\hat{f} - \hat{P}_2)^2 \hat{V}(\hat{R}) + 2\hat{R}\hat{P}_1^2 (\hat{f} - \hat{P}_2) \widehat{\mathrm{Cov}}(\hat{f}, \hat{R})}{(\hat{P}_1 - \hat{P}_2)^2} - \frac{2\hat{P}_1 \hat{P}_2 \hat{N}_1 \hat{N}_2 \widehat{\mathrm{Cov}}(\hat{P}_1, \hat{P}_2)}{(\hat{P}_1 - \hat{P}_2)^2}$$

where $\hat{N}_2 = \hat{N}_{\mathrm{cir}} - \hat{R}$. In eqs. 13 and 14, the first term is the uncertainty in the survey results, the middle term is the uncertainty in the removals, and the third term is the reduction in variance due to the reoccupying of the same stations in survey 2 as in survey 1.

If the catchability of the two types of animals in the surveys is equal (so that the change-in-ratio estimate of total population, i.e., x- and y-types, is valid), then the variance of the estimated total can be estimated as

$$[15]\quad \hat{V}(\hat{N}_{\text{cir}}) = \frac{\hat{N}_{\text{cir}}^2\,\hat{V}(\hat{P}_1) + \hat{N}_2^2\,\hat{V}(\hat{P}_2)}{(\hat{P}_1 - \hat{P}_2)^2} + \frac{\hat{R}^2\,\hat{V}(\hat{f}) + (\hat{f} - \hat{P}_2)^2\,\hat{V}(\hat{R}) + 2\hat{R}\,(\hat{f} - \hat{P}_2)\,\widehat{\text{Cov}}(\hat{f}, \hat{R})}{(\hat{P}_1 - \hat{P}_2)^2} - \frac{2\hat{N}_{\text{cir}}\,\hat{N}_2\,\widehat{\text{Cov}}(\hat{P}_1, \hat{P}_2)}{(\hat{P}_1 - \hat{P}_2)^2}\,.$$

Also, if the catchability of the two types of animals in the surveys is equal then the total population of the two types can be estimated by $\hat{N}_{\text{ir}} = \hat{X}_{\text{ir}}/\hat{P}_1$, with variance estimated by

$$[16]\quad \hat{V}(\hat{N}_{\text{ir}}) = \frac{\hat{N}_{\text{ir}}^2\,(\hat{V}(\hat{c}_{x1}) + \hat{V}(\hat{c}_{x2})) + \hat{c}_{x1}^2\hat{R}_x^2\,\hat{P}_1^{-4}\,\hat{V}(\hat{P}_1)}{(\hat{c}_{x1} - \hat{c}_{x2})^2} + \frac{\hat{c}_{x1}^2\,\hat{V}(\hat{R}_x)}{\hat{P}_1^2\,(\hat{c}_{x1} - \hat{c}_{x2})^2} - \frac{2\hat{N}_{\text{ir}}^2\,\widehat{\text{Cov}}(\hat{c}_{x1}, \hat{c}_{x2})}{(\hat{c}_{x1} - \hat{c}_{x2})^2}\,.$$

We estimate the total population size this way, rather than as the sum of two index-removal estimates, because one of the classes is unexploited (or lightly exploited) thus giving rise to no estimate (or a very imprecise estimate) for the class.

Variance formulae for estimates of exploitation rate and catchability coefficient are given in Appendix 1.

It is also possible to use the Taylor's series approach to estimate the bias of the estimated population parameters (Appendix 2).

Evaluation of the efficient design

Modeling catches

We conducted a simulation study to see how the change-in-ratio and index-removal methods perform when animals have a patchy distribution with a degree of persistence in spatial pattern over time. This is a common pattern in nature and has been observed for snow crab as shown in the example in the next section.

When animals are independently and randomly distributed over space, the catch obtained from a standard unit of sampling effort has a Poisson distribution and the variance of catches is equal to the mean catch. Also, the sum of the catches from J units of sampling effort (i.e., from J locations) has a Poisson distribution with parameter equal to J times the mean catch per unit effort. Thus, following Chapman and Murphy (1965), we might assume that the catches of x- and y-type animals in the surveys have the following distributions:

$$x_{1j} \sim \text{Poisson}(q_1 f_1 X_1), j = 1, 2, \ldots J$$
$$y_{1j} \sim \text{Poisson}(q_1 f_1 Y_1), j = 1, 2, \ldots J$$
$$x_{2j} \sim \text{Poisson}(q_2 f_2 (X_1 - R_x)), j = 1, 2, \ldots J$$
$$y_{2j} \sim \text{Poisson}(q_2 f_2 (Y_1 - R_y)), j = 1, 2, \ldots J$$

where q_i is the catchability coefficient in the i^{th} survey, f_i is the sampling effort expended at each location in the i^{th} survey, and the other symbols are as defined previously. For index-removal estimation, we assume further that $q_1 = q_2$. We refer to this as the common Poisson model because the catches for each location follow a common Poisson distribution (there is no location effect). It is clear that, because there is no location effect, there is no advantage to reoccupying the same stations in the second survey as in the first.

In the common Poisson model, the variances of the catches equal the means. However, when animals have a patchy distribution the variance is greater than the mean. We therefore modify the common Poisson model to allow for this overdispersion by introducing random location effects. (There may be a spatial correlation in animal distribution but, because the sampling station locations are chosen randomly, the station effects are independent, identically distributed random variables.) For convenience, we let the location effects have a gamma distribution. The gamma distribution is extremely flexible and, when a variable Z follows a Poisson distribution whose parameter, λ, is drawn from a gamma distribution, the unconditional distribution of Z is negative binomial.

We used two models to simulate sampling from a population. In model I, the random location effects for x-type animals are independent of those for y-type animals. The sample catches then have the following distributions:

$$x_{1j} \sim \text{Poisson}(q_1 f_1 X_1\,\delta x_{1j}), j = 1, 2, \ldots J$$
$$y_{1j} \sim \text{Poisson}(q_1 f_1 Y_1\,\delta y_{1j}), j = 1, 2, \ldots J$$
$$x_{2j} \sim \text{Poisson}(q_2 f_2 (X_1 - R_x)\,\delta x_{2j}), j = 1, 2, \ldots J$$
$$y_{2j} \sim \text{Poisson}(q_2 f_2 (Y_1 - R_y)\,\delta y_{2j}), j = 1, 2, \ldots J$$

where

$$\delta x_{ij} \sim \text{Gamma}(\alpha_1, \beta_1), i = 1, 2; j = 1, 2, \ldots J$$
$$\delta y_{ij} \sim \text{Gamma}(\alpha_2, \beta_2), i = 1, 2; j = 1, 2, \ldots J.$$

The expected value of the catch of x-type animals (at any location) is

$$E(x_{ij}) = q_i f_i X_i \alpha_1 \beta_1$$

while for y-type animals it is

$$E(y_{ij}) = q_i f_i Y_i \alpha_2 \beta_2.$$

In order to keep the catchabilities for the two types of animals equal, we introduce the constraint that $\alpha_1\beta_1 = \alpha_2\beta_2$. Note that for the paired design, $\delta x_{1j} = \delta x_{2j}$ and $\delta y_{1j} = \delta y_{2j}$ because the same stations are used for the two surveys. For index-removal estimation, q_1 is set equal to q_2.

In general, the catches of the two subclasses at a given time are correlated and the catches of a given subclass (at a set of fixed locations) are correlated over time. We use model II, a generalization of model I, to mimic this. The catches are distributed as follows:

$$x_{1j} \sim \text{Poisson}(q_1 f_1 X_1 (\delta_{1j} + \delta x_{1j})), j = 1, 2, \ldots J$$
$$y_{1j} \sim \text{Poisson}(q_1 f_1 Y_1 (\delta_{1j} + \delta y_{1j})), j = 1, 2, \ldots J$$
$$x_{2j} \sim \text{Poisson}(q_2 f_2 (X_1 - R_x) (\delta_{2j} + \delta x_{2j})), j = 1, 2, \ldots J$$
$$y_{2j} \sim \text{Poisson}(q_2 f_2 (Y_1 - R_y) (\delta_{2j} + \delta y_{2j})), j = 1, 2, \ldots J$$

where

$$\delta_{ij} \sim \text{Gamma}(\alpha, \beta), i = 1, 2; j = 1, 2, \ldots J$$
$$\delta x_{ij} \sim \text{Gamma}(\alpha_1, \beta_1), i = 1, 2; j = 1, 2, \ldots J$$
$$\delta y_{ij} \sim \text{Gamma}(\alpha_2, \beta_2), i = 1, 2; j = 1, 2, \ldots J.$$

Again, we introduce the constraint that $\alpha_1\beta_1 = \alpha_2\beta_2$ so that the

Table 1. Comparison of change-in-ratio estimates for the paired and unpaired designs when data are generated according to overdispersion model I.

	Bias($\hat{X}_{cir}$)		SE($\hat{X}_{cir}$)	
n_i/N_i	Paired	Unpaired	Paired	Unpaired
0.9	2.0	49.6	42.3	2 079.6
0.7	2.8	59.3	48.1	692.9
0.5	4.4	70.4	57.3	2 245.3
0.3	7.8	77.0	75.5	2 323.1
0.1	31.6	78.5	173.2	8 659.3
	Bias($\hat{Y}_{cir}$)		SE($\hat{Y}_{cir}$)	
0.9	26.5	98.1	235.2	3 508.1
0.7	27.9	115.5	239.2	1 141.9
0.5	29.7	135.0	248.5	3 828.0
0.3	38.1	146.1	269.2	4 041.4
0.1	81.7	158.9	405.1	11 089.9

Note: Estimates are based on 40 000 simulations of the paired design and 100 000 simulations for the unpaired design. R_x/X_1 (fraction of the x-type population removed) was fixed at 0.7, R_y/Y_1 was fixed at 10%. n_i/N_i is the expected fraction of the population seen in each survey. Other parameters are given in the text. The sizes of the x- and y-populations being estimated are 1000.

Table 2. Comparison of change-in-ratio estimates for the paired and unpaired designs when data are generated according to overdispersion model II.

	Bias($\hat{X}_{cir}$)		SE($\hat{X}_{cir}$)	
n_i/N_i	Paired	Unpaired	Paired	Unpaired
0.9	2.8	26.1	41.7	132.5
0.7	2.9	27.0	47.4	165.3
0.5	4.7	30.8	56.6	126.8
0.3	8.1	34.8	75.6	288.2
0.1	30.8	64.4	186.2	1014.6
	Bias($\hat{Y}_{cir}$)		SE($\hat{Y}_{cir}$)	
0.9	15.7	52.0	172.2	285.9
0.7	16.0	52.4	180.1	325.6
0.5	19.6	58.3	188.6	372.1
0.3	25.8	66.0	214.1	462.0
0.1	67.5	119.5	370.8	1693.7

Note: Estimates are based on 40 000 simulations. R_x/X_1 (fraction of the x-type population removed) was fixed at 0.7, R_y/Y_1 was fixed at 10%. n_i/N_i is the expected fraction of the population seen in each survey. Other parameters are given in the text. The sizes of the x- and y-populations being estimated are 1000.

catchabilities for the two types are equal. Also, note that for the paired design $\delta_{1j} = \delta_{2j}$, $\delta x_{1j} = \delta x_{2j}$, and $\delta y_{1j} = \delta y_{2j}$.

Design of the simulation study

We compared the bias and standard error of the estimates for the paired and unpaired design. The simulation data were generated from the two overdispersion models under the following conditions: (*i*) population size $N_1 = 2000$ with $X_1 = 1000$ and $Y_1 = 1000$; (*ii*) $q_1 = q_2 = 0.001$ with effort chosen so that the expected fraction of the population seen in each survey (n_i/N_i) ranged from 0.1 to 0.9 by increments of 0.2; (*iii*) 10% of y-type animals removed (assumed known exactly); (*iv*) 70% of x-type animals removed; (*v*) overdispersion parameters $\alpha_1 = \alpha_2 = 1.0$, $\beta_1 = \beta_2 = 1.0$ for model I and $\alpha = \alpha_1 = \alpha_2 = 0.5$, $\beta = \beta_1 = \beta_2 = 1.0$ for model II (thus, expected catches in the surveys are the same for the two models); (*vi*) at least 10 000 data sets were generated for each scenario; and (*vii*) new values of the δ_{ij}, δx_{ij}, and δy_{ij} were generated for each simulated data set. These are conditions in which the methods can be expected to perform well. We also conducted an additional simulation which may be more typical of a commercial fishery. The initial population size of x-type animals was 3.2×10^6 animals while for y-type animals it was 3.5×10^6. The exploitation rates on the x- and y-type animals were 18 and 5%, respectively. The expected fraction of the population seen in each survey was 1%. The overdispersion parameters were $\alpha = \alpha_1 = \alpha_2 = 0.5$ and $\beta = \beta_1 = \beta_2 = 1.0$. The simulation of additional scenarios is discussed by Chen (1995).

Whenever an estimate was infeasible (e.g., negative) we replaced it with the minimum feasible estimate which is equal to the removal plus the number of animals seen in the second survey.

Table 3. Comparison of unpaired and paired designs for change-in-ratio and index-removal estimation for a simulated population with exploitation rates of 18 and 5% for x- and y-type animals, respectively.

Type	$E(\hat{X}_1)$	SE($\hat{X}_1$)	$E(\hat{Y}_1)$	SE($\hat{Y}_1$)
Change-in-ratio estimator				
Unpaired, no lower bound	1.31	2080	1.66	2380
Unpaired, lower bound	10.7	1380	12.4	1710
Paired	3.21	0.273	3.57	0.728
Index-removal estimator				
Paired[a]	3.21	0.123	—	—

Note: True population sizes were $X_1 = 3.2 \times 10^6$ and $Y_1 = 3.5 \times 10^6$. Number of simulated data sets was 5 000 000 for the change-in-ratio estimators and 10 000 for the index-removal methods. See text for other parameters. When estimated population size was infeasible (e.g., negative) the estimate was replaced by the minimal estimate (lower bound) equal to the removal plus the number of animals observed in the second survey. $E(\cdot)$ = expected value; SE($\cdot$) = standard error. Estimates should be multiplied by 10^6.

[a]Standard errors for the unpaired cases (with and without a lower bound) are in excess of 100×10^6 but estimates are not precise and are not reported here.

Simulation results

Change-in-ratio estimation

In all of the simulations the paired design worked far better than the unpaired design (Tables 1–3). In fact, the variability in the results with the unpaired design was so large that 10 000 simulations were not enough to enable us to determine precisely the bias and variance. Nonetheless, it is clear that even when 70% of the x-type population is removed, compared to only 10% of the y-type population, the bias is substantial and the standard errors are so large under the unpaired design as to call the utility of the estimator into question.

The simulation of conditions representing a fishery with an

Table 4. Comparison of index-removal estimates for the paired and unpaired designs when data are generated according to overdispersion model I.

	Bias($\hat{X}_{ir}$)		SE($\hat{X}_{ir}$)	
x_i/X_i	Paired	Unpaired	Paired	Unpaired
0.9	1.7	19.8	30.3	107.8
0.7	1.9	21.4	34.4	114.6
0.5	2.6	21.9	41.0	115.6
0.3	4.9	24.9	54.2	127.6
0.1	14.3	35.2	101.0	228.2

Note: Estimates are based on 10 000 simulations. R_x/X_1 is the fraction of the x-type population removed and x_i/X_i is the fraction of the x-type population seen in each survey. Other parameters are given in the text. The size of the x-population being estimated is 1000.

Table 5. Results from the pre- and post-season surveys of snow crab in St. Mary's Bay using small mesh traps.

Catch		Catch rate	Covariance	Correlation, r
X before	X before	44.0727	11.3911	1.0
X before	Y before		14.0993	0.42776
Y before	Y before	100.3455	95.3727	1.0
X before	X after		5.3555	0.58060
X after	Y before		6.8309	0.25593
X after	X after	33.0727	7.4693	1.0
X before	Y after		14.0305	0.44432
Y before	Y after		66.9709	0.73295
X after	Y after		9.8238	0.38419
Y after	Y after	88.1091	87.5364	1.0

$\hat{P}_1 = 0.30517$; $\hat{V}(\hat{P}_1) = 0.40287 \times 10^{-3}$; $\widehat{\text{Cov}}(\hat{P}_1, \hat{P}_2) = 0.23468 \times 10^{-3}$; $\hat{P}_2 = 0.27292$; and $\hat{V}(\hat{P}_2) = 0.44739 \times 10^{-3}$.

Note: Covariances are covariances of the means, i.e., covariance of the paired observations divided by the number of sampling stations (= 55). (X-type = legal size, Y-type = sublegal size.)

exploitation rate of 18% on x-type animals and 5% on y-type animals was very informative. The results for the unpaired design were so enormously variable that it took 1 880 000 simulated data sets to get the estimates of expected value and standard error to stabilize somewhat but even after this many simulations the results were not very precise. For this scenario, the change-in-ratio estimator is badly biased and has enormous variance when the unpaired design is used (Table 3). Also, infeasible (negative) estimates arise frequently. In contrast, the estimator worked much better with the paired design: none of the estimates were infeasible, the bias appeared small, but the standard errors were very large.

Index-removal estimation

Under model I, the paired design provided nearly unbiased estimates and small standard errors for all scenarios considered (Table 4). The unpaired results were uniformly inferior to the paired results.

For the exploited fishery scenario (Table 3), the results are similar to those found for the change-in-ratio estimator: the unpaired design performed very poorly but the paired design always produced feasible estimates, appeared to have minimal bias, and had a very small standard error.

Table 6. Results from the pre- and post-season surveys of snow crab in St. Mary's Bay using large mesh traps.

Catch		Catch rate	Covariance	Correlation, r
X before	X before	61.3454	17.5173	1.0
X before	Y before		15.0271	0.53503
Y before	Y before	67.4546	45.0327	1.0
X before	X after		10.7164	0.72335
X after	Y before		9.1520	0.38530
X after	X after	44.4546	12.5291	1.0
X before	Y after		12.5767	0.46584
Y before	Y after		31.4836	0.72732
X after	Y after		13.4195	0.58773
Y after	Y after	58.8000	41.6091	1.0

$\hat{P}_1 = 0.47628$; $\hat{V}(\hat{P}_1) = 0.45351 \times 10^{-3}$, $\widehat{\text{Cov}}(\hat{P}_1, \hat{P}_2) = 0.32587 \times 10^{-3}$; $\hat{P}_2 = 0.43053$; and $\hat{V}(\hat{P}_2) = 0.48732 \times 10^{-3}$.

Note: Covariances are covariances of the means, i.e., covariance of the paired observations divided by the number of sampling stations (= 55). (X-type = legal size, Y-type = sublegal size.)

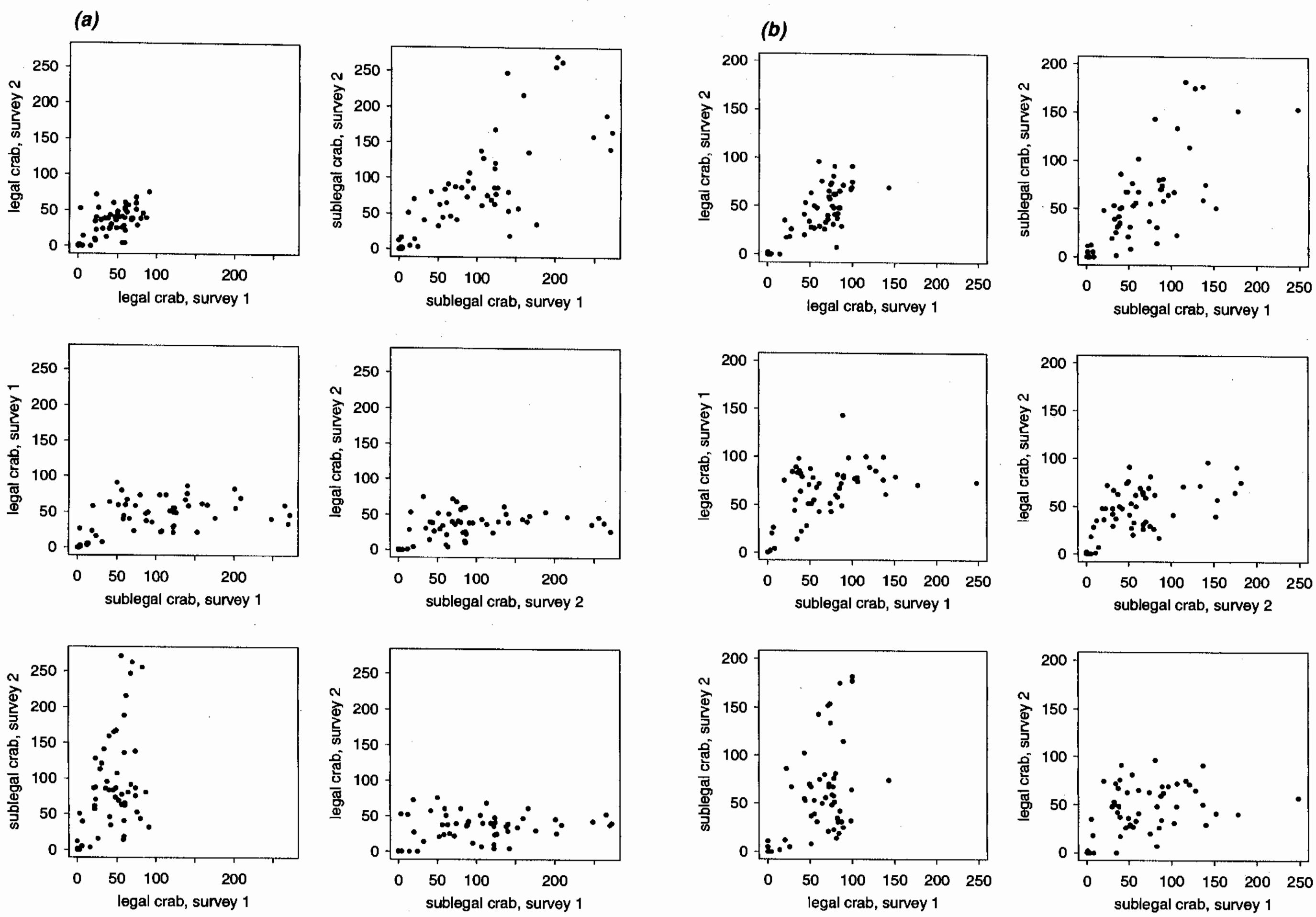
Example: snow crab fishery in St. Mary's Bay, Newfoundland

The fisheries for snow crab in Atlantic Canada are exclusively trap fisheries and only males above the minimum legal size of 95 mm carapace width are harvested. There is a mesh size restriction to reduce the catch of sublegal-sized crab and fishermen have been taught to return sublegal-sized crab to the water as quickly as possible to protect the prerecruits. We define x-type animals to be males ≥95 mm in width and y-type animals to be males with width ≥78 but <95 mm. In 1992, the fishery in St. Mary's Bay occurred from September 1 to September 10. From sales slip data, it is known that 306 t of snow crab were harvested. Extensive sampling of crab in the fish plants revealed that the mean weight of individual crab was 480.49 g and that 90.826% of the harvested crab were of legal size (x-type). Thus, we calculate that 578 425 legal-sized and 58 425 sublegal-sized crab were harvested. For purposes of this analysis we assume these figures are exact.

The preseason research survey was conducted from August 20 to August 28 and the postseason survey was from September 14 to September 28. Fifty-five fleets of traps were set at randomly selected locations in that part of the bay deeper than 40 m (approximately 660 km^2). The same locations were occupied in both surveys. Each fleet of traps contained 2 large-meshed and 2 small-meshed traps. The large-meshed traps were the same type as used in the commercial fishery. The primary sampling unit was the combined catch from all of the traps of a given type in a fleet of traps. However, catches were expressed on a per-trap basis to account for the fact that occasionally a trap in a fleet didn't fish properly. The survey methods are described in Dawe et al. (1993).

For both small- and large-meshed traps, the proportion of legal-sized crab in the second survey was slightly lower than in the first survey (Tables 5 and 6). For small-meshed traps, the proportion declined from 0.31 to 0.27 between the two

Fig. 1. Scatterplots of catches of crab of two size-classes in survey i versus catches of crab at the same location in survey j for $i, j = 1, 2$. Note that the catches of legal-sized crab in survey 1 are highly correlated with catches of legal-sized crab in survey 2, and similarly for sublegal-sized crab. In contrast, catches of legal-sized crab are only weakly correlated with catches of sublegal-sized crab both within a survey and between surveys. (*a*) Catches from small-mesh traps. (*b*) Catches from large-mesh traps.



surveys while for large-meshed traps the proportion declined from 0.48 to 0.43. The catch rates also declined slightly from the first survey to the second (Tables 5 and 6). For small-meshed traps, the catch rate declined from 44 to 33 legal-sized crab per trap and from 100 to 88 sublegal-sized crab per trap. For large mesh traps, the catch rate declined from 61 to 44 legal-sized crab per trap and from 67 to 59 sublegal-sized crab per trap.

The correlation, r, between catch per trap in the preseason survey and the catch at the corresponding location in the postseason survey was large and positive (Tables 5 and 6, Fig. 1) thus indicating that increased efficiency can indeed be attained by reusing rather than rerandomizing the sampling locations. For small-mesh traps, the correlation between catches of legal-sized crab in the two surveys was 0.58 and for sublegal-sized crab it was 0.73. For large-mesh traps, the correlation was 0.72 for legal-sized crab and 0.73 for sublegal-size crab.

The estimates of the number of legal-sized crab at the beginning of the season ranged from 2.1 to 3.8 million animals with standard errors for the paired design ranging from 0.3 million (for the smallest estimate) to 2.1 million (for the largest estimate) (Table 7). The standard errors were also calculated without including the covariance term. The reduction in estimated standard error achieved by including the covariance term was 33–47%.

The exploitation rate was estimated to be from 15 to 28% (Table 8) with standard errors ranging from 4 to 8% for the paired design. As with the estimates of population size, these values were 33–47% less than those obtained when the covariance term was ignored.

Incorporating information about relative catchability in index-removal estimators

The problem of heterogeneity of capture probabilities can be minimized by making separate estimates for various subsets of the population. For example, separate estimates could be made for males, and for females, or for different size groups of animals. However, when information is available on the relative catchability of different groups, this information can be incorporated in the estimation procedure to increase the statistical efficiency of the estimator. For example, one may have good

Table 7. Estimates of initial population of legal-sized snow crab ($\times 10^6$).

Method	Estimate	SE with covariance	SE without covariance	% reduction in SE
Estimates from large-mesh traps				
cir	3.1673	1.0642	1.9199	45
ir	2.1008	0.3096	0.5797	47
Estimates from small-mesh traps				
cir	3.8287	2.1465	3.1974	33
ir	2.3175	0.5094	0.7850	35

Note: To estimate how reoccupying stations improves efficiency, the standard error (SE) was calculated with and without the covariance term.

Table 8. Estimates of exploitation rate for legal-sized snow crab in St. Mary's Bay.

Method	Estimate	SE with covariance	SE without covariance	% reduction in SE
Estimates from large-mesh traps				
cir	0.183	0.0614	0.1107	45
ir	0.275	0.0406	0.0760	47
Estimates from small-mesh traps				
cir	0.151	0.0847	0.1262	33
ir	0.250	0.0549	0.0845	35

reason to believe that male crab are more catchable than females in a trap survey because of differential behavior or differential body size among the sexes. Similarly, one may believe that large crab may be more catchable by traps than small crab. In these cases, one may wish to introduce order restrictions in the estimation procedure to ensure that the estimated catchabilities are consistent with the available information on relative catchabilities of the various groups.

We consider a suite of four models that vary in the amount of information assumed about the relative catchabilities of the different groups. The simplest approach is to make separate, independent estimates for each group. If qualitative information is available about the relative catchabilities of the groups, then one can introduce order restrictions for the catchability coefficients. One might also assume a functional relationship for the way catchability coefficients vary with a covariate. In particular, the catchability coefficient might be a logistic function of body size. In this case, we would estimate the parameters of the functional relationship rather than the catchability coefficients for each size group. Finally, if a sampling gear selectivity curve is available from some other study, then the parameter estimates of the selectivity curve can be incorporated directly in the population estimation procedure.

We illustrate the approach by assuming that the catches per unit of sampling effort in the surveys follow Poisson or multinomial distributions. This is consistent with previous treatments of the subject in the literature. It is possible to assume catch rate follows a different distribution such as a normal distribution as suggested by Routledge (1989). The normal distribution appears to be a more reasonable model for many fishery applications and the modifications to handle this are straightforward.

Four models for catches that follow Poisson distributions

We assume that the expected value of the total number of animals caught when one unit of sampling effort is expended at each of n randomly selected locations is given by $E(C) = qnX = \lambda$ (say), where $E(\cdot)$ denotes expected value of the quantity in parentheses, q is the catchability coefficient, n is the number of units of sampling effort (locations), and X is the population size. Thus, catch is assumed proportional to sampling effort and to abundance. This assumption is justified if the sampling is with replacement (animals are released unharmed after being caught) or the fraction of the population caught is negligible so that the population size, X, does not change due to the random sampling. Furthermore, we assume that the total number of animals caught during the survey, C, follows a Poisson distribution with parameter λ, i.e., $C \sim P(\lambda)$. Thus, the probability density function for the number of animals caught is

$$f(C) = \frac{\lambda^C e^{-\lambda}}{C!} = \frac{(qnX)^C e^{-qnX}}{C!} .$$

If R animals are removed from the population, the abundance becomes $X - R$ and, under the assumption that expected value of the catch is proportional to abundance, the expected value of the catch becomes $E(C) = qn(X - R)$. If we assume that the distribution of catch remains Poisson then we can write the likelihood, Λ, for obtaining a series of catches $\{C_1, C_2, \ldots C_J\}$ from J surveys having respective sampling efforts $\{n_1, n_2, \ldots n_J\}$ as

$$[18] \quad \Lambda = \prod_{j=1}^{J} \frac{\lambda_j^{C_j} e^{-\lambda_j}}{C_j!} = \prod_{j=1}^{J} \frac{(qn_j X_j)^{C_j} e^{-qn_j X_j}}{C_j!} .$$

where λ_j is the Poisson parameter for the j^{th} random survey and X_j is the number of animals in the population just before the j^{th} survey. Thus, X_1 is equal to the original population X, and

$$[19] \quad X_j = X - \sum_{k=1}^{j-1} R_j \quad \text{for } j \geq 2$$

where R_k is the number of animals removed from the population after the k^{th} survey. Here, we have treated the removals R_k as known, fixed values.

There are two unknowns, q and X_1, in eq. 18. When the data consist of two surveys and one removal, the estimates which maximize the likelihood are given by

$$\hat{X}_1 = \frac{\hat{c}_1 \hat{R}}{\hat{c}_1 - \hat{c}_2}$$

$$\hat{q} = \frac{\hat{c}_1}{\hat{X}_1} = \frac{\hat{c}_1 - \hat{c}_2}{\hat{R}} .$$

Here, $\hat{c}_j$ is the catch per unit of sampling effort in the j^{th} survey for $j = 1, 2$.

Method 1: independent estimates by size group

Suppose we have reason to believe that the I subgroups in the population have different catchabilities, q_i. Suppose, further, that we believe the catches of the various subgroups are independent Poisson random variables. This assumption is made explicitly for some change-in-ratio estimation models (see Seber 1982). Then the likelihood (eq. 18) can be generalized by the introduction of an index i denoting subgroup-specific population sizes and catchabilities. Thus, the likelihood becomes

$$[20] \quad \Lambda = \prod_{i=1}^{I} \prod_{j=1}^{J} \frac{\lambda_{ij}^{\ C_{ij}} e^{-\lambda_{ij}}}{C_{ij}\,!} = \frac{(q_i\, n_j\, X_{ij})^{C_{ij}}\, e^{-q_i n_j X_{ij}}}{C_{ij}\,!} .$$

Here, the X_{ij} are defined in a manner analogous to eq. 19. That is, X_{i1} is the original number of animals in the population in subgroup i and

$$[21] \quad X_{ij} = X_{i1} - \sum_{k=1}^{j-1} R_{ik} \ \text{for}\ j \geq 2$$

where R_{ik} is the number of animals removed from the i^{th} subgroup of the population after the k^{th} survey. Here, we have again treated the removals R_{ik} as known, fixed values.

Equation 20 has $2I$ unknowns: I initial abundances and I catchability coefficients. It is easily verified that maximizing eq. 20 with respect to the $2I$ unknowns is equivalent to maximizing eq. 18 separately for each subgroup in the population.

Method 2: introducing order restrictions for the catchabilities

Suppose we have good reason to believe that $q_1 < q_2 < \ldots q_I$ and we wish to introduce these order restrictions into the estimation procedure. Let $q_0 = 0$ and let $q_i = q_{i-1} + \delta_i^2$, for $i = 1, 2, .. I$, and substitute these definitions into the likelihood (eq. 20). We now have I initial abundances and I values of δ_i^2 to estimate. Note that, regardless of the value of the estimate $\hat{\delta}_i$, the value of $\hat{\delta}_i^2$ must be nonnegative and, thus, the estimate of q_i must be greater than or equal to the estimate of q_{i-1}.

Method 3: introducing a functional relationship for catchability

Often, the catchability of an animal will vary with the animal's body size. For example, the chances of a fish escaping through the meshes of a trawl generally decrease as the size of the fish increases. In contrast, the ability of some animals to avoid sampling gear may increase as the animals increase in age or size. In fisheries work, it is common to model the selectivity of fishing gear as a logistic function of body size. Thus, the proportion of the animals of length l that is retained by the gear, $p(l)$, can be described by

$$[22] \quad p(l) = \frac{1}{1 + \exp\left(-\alpha\,(1 - l_0)\right)}$$

where α is a shape parameter and l_0 is a location parameter. The catchability of animals in the i^{th} size group would then be proportional to the selectivity for that group:

$$[23] \quad q_i = \beta\, p(l)$$

where β is an additional parameter relating the catchability to the selectivity of the gear. Equation 23 can be substituted for the q_i in eq. 20. In this case, we estimate the parameters α and l_0 of the logistic curve and the scaling parameter β, instead of the I catchabilities, q_i. When only a limited range of sizes of animals is caught, it might be better to model catchability as an asymptotic or linear function of size. The basic idea remains the same.

Method 4: using independent estimates of gear selectivity

It is often the case that the gear selectivity can be estimated by comparing the catches from two sampling gears with different mesh sizes. If gear selectivity parameter estimates are available from an independent study then one need only estimate the initial population sizes, X_{i1}, and the scaling parameter, β, of the logistic curve.

Discussion

Efficient design

The present study appears to be the first to point out the potential gains in efficiency obtainable by the reuse of the same stations in the two surveys. These gains can be achieved at no additional cost of sampling, i.e., without increasing the sampling effort. In the snow crab example, the increase in precision was substantial: standard errors of estimated population size, calculated with the covariance term, were 33–47% less than those calculated without the covariance term. In the simulation studies, the variances of the catch rates were considerably smaller than those observed in St. Mary's Bay. However, the correlations (between types of animals in a survey and between surveys) were grossly similar to what was observed for St. Mary's Bay (see Chen 1995, Table 11). The simulations clearly demonstrate the reduction in bias and increase in efficiency attainable through use of the efficient design.

Gains in efficiency can be expected when the spatial pattern persists from one survey to the next and when the scale of the patchiness is large enough that the researcher is able to place the sampling gear in the same patches in the subsequent survey. Short-term movements and migrations and imprecision in setting fishing gear will lessen the effectiveness of reusing sampling locations.

Use of the paired design also results in a substantial reduction in the small-sample bias and in the incidence of infeasible estimates when sample sizes are small. Bias of the change-in-ratio and index-removal estimators has not previously been studied although it is known that the estimators are asymptotically (as sample sizes approach infinity) unbiased. Bias of the estimates of the total and x-type population sizes, when the covariance (between surveys) is zero, is positive for both the change-in-ratio and index-removal methods. Magnitude of bias increases as the change in catch rates approaches zero and as the change in proportions approaches zero. Positive bias in estimates of population size implies that the estimates of exploitation rate have a negative bias.

Change-in-ratio estimators can also be used to estimate the relative survival rate of two groups (see Seber 1982). For example, Frusher et al. (1998) estimate the ratio of survival of legal lobsters to that of sublegal-sized lobsters of the same sex. By similar methods to those used here, it can be shown that gains in efficiency can be achieved for this estimator when the sampling stations in the second survey are the same as in the first survey. Gains in efficiency with the paired design can also be expected for the change-in-ratio estimator of relative habitat usage and relative catchability described by Heimbuch and Hoenig (1989).

Incorporating information on catchability

The methods described for incorporating information on relative catchability of different population components can be used as part of a general model building strategy. One can start by looking at separate estimates of catchability by size group. If these show a general trend or pattern that is consistent with expectation based on knowledge of the biology of the species and the characteristics of the sampling gear, then the estimates might be smoothed somewhat by imposing order restrictions. One might also estimate the parameters of a selectivity-with-size model. However, this would require sufficient contrast in the data, i.e., a sufficient range of sizes in the data. One could also use assumed selectivity parameters if these are available from an external study. A likelihood ratio test could be used to test if selectivity parameters estimated by the index-removal method are significantly different from assumed values from an external study.

The problem of heterogeneity of capture probabilities also occurs with removal estimators. One way to handle this is to make separate estimates for each component of the population. In this case, one may wish to constrain or model the catchability coefficients as functions of the sizes of the animals as was done for the index-removal estimator.

Conclusions about snow crab in St. Mary's Bay

The estimates of population size are rather variable and the standard errors are quite large. This is what would be expected if the removal accounted for only a small fraction of the legal-sized population. Judging from the small change in catch rate and the small change in composition, this appears to be the case. Estimates of the fraction of the legal-sized population removed vary from 15 to 28% (Table 8). Thus, no matter which portion of the data is considered, and no matter which analytical method is used, the results are that the population in St. Mary's Bay is lightly exploited relative to other areas where target exploitation rates of 50–60% have been achieved. These results are consistent with estimates for St. Mary's Bay for the previous year (1991) of 9–29% exploitation rate (Dawe et al. 1993).

We computed index-removal estimates of population size of legal-sized crab by 5-mm length-class. We expected estimates of the catchability coefficient to show an increasing trend with size but, instead, the estimates decreased with size. We have no explanation for this.

The mean catch rate for sublegal-sized crab declined from the first survey to the second one. Although the decline is not statistically significant, it is interesting to consider the implications of such a decline. One possibility is that the catchability of all crab declined over the course of the study. In this case, the estimated decline in catch rate for legal-sized crab would be too large and the index-removal estimate of population size of legal-sized crab would be too low. The change-in-ratio estimator would be unaffected by a change in catchability provided that both size groups were equally affected (same percentage decline in catchability). Another possibility is that the population of sublegal-sized crab declined not only because of the known harvest but also because of mortality experienced by discarded sublegal-sized crab. In this case, the index-removal estimator of the legal-sized crab population is unaffected but the change-in-ratio estimator is biased low (see Seber 1982, page 361). The fact that the interval of time between the start of the first survey and the end of the second was short (40 days) suggests that any change in catchability (as well as failure of the assumption of closure) should be minimal. Studies have shown that if snow crab are returned to the water quickly the probability of survival can be very high and these studies have been shown to crab fishermen who now appear to recognize that discarded small crab are their future income. Therefore, it is suspected that discard mortality is relatively small.

Estimates of population size from large-mesh trap data are smaller than the corresponding estimate from small-mesh trap data for both the change-in-ratio and index-removal methods. However, the reverse was observed in the previous year's study (Dawe et al. 1993) so this does not appear to be significant.

It appears that change-in-ratio and index-removal estimators may be quite useful for assessing invertebrate stocks and estimating a variety of parameters. However, up until now it appeared to be difficult to obtain precise estimates. The current work suggests that through careful study design it may be possible to achieve acceptable precision.

Acknowlegments

We thank Paul Beck and Joe Drew for help with all aspects of the field work. Robert Kennedy and Mark Udevitz provided helpful comments.

References

Chapman, D.G., and Murphy, G.I. 1965. Estimates of mortality and population from survey-removal records. Biometrics, **21**: 921–935.

Chen, C.L. 1995. Change-in-ratio methods. Ph.D. thesis. Department of Statistics, North Carolina State University, Raleigh, N.C.

Cochran, W.G. 1977. Sampling techniques. Third edition. John Wiley & Sons, Inc., New York, N.Y.

Dawe, E.G., Hoenig, J.M., and Xu, X. 1993. Change-in-ratio and index-removal methods for population assessment and their application to snow crab (*Chionoecetes opilio*). Can. J. Fish. Aquat. Sci. **50**: 1467–1476.

DeLury, D.B. 1947. On the estimation of biological populations. Biometrics, **3**: 145–167.

Eberhardt, L.L. 1982. Calibrating an index using removal data. J. Wildl. Manage. **46**: 734–740.

Frusher, S.D., Kennedy, R.B., and Gibson, I.D. 1998. Preliminary estimates of exploitation rates in the Tasmanian rock lobster (*Jasus edwardsii*) fishery using the change-in-ratio and index-removal techniques with tag–recapture data. *In* Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. *Edited by* G.S. Jamieson and A. Campbell. Can. Spec. Publ. Fish. Aquat. Sci. 125. pp. 63–71.

Heimbuch, D.G., and Hoenig, J.M. 1989. Change-in-ratio estimators for habitat usage and relative population size. Biometrics, **45**: 439–451.

Hoenig, J.M., and Pollock, K.H. 1998. Index-removal methods. *In* Encylopedia of statistical sciences. Update volume. *Edited by* S. Kotz, C.B. Read, and D.L. Banks. John Wiley & Sons, Inc., New York, N.Y. pp. 341–346.

Kelker, G.H. 1940. Estimating deer populations by a differential hunting loss in the sexes. Proc. Utah Acad. Sci. Arts Lett. **17**: 6–69.

Leslie, P.H., and Davis, D.H.S. 1939. An attempt to determine the absolute number of rats on a given area. J. Anim. Ecol. **8**: 94–113.

Petrides, G.A. 1949. View points on the analysis of open season sex

and age ratios. Trans. N. Am. Wildl. Nat. Resour. Conf. **14**: 391–410.
Pollock, K.H., and Hoenig, J.M. 1998. Change in ratio estimators. *In* Encylopedia of statistical sciences. Update volume. *Edited by* S. Kotz, C.B. Read, and D.L. Banks. John Wiley & Sons, Inc., New York, N.Y. pp. 109–112.
Pollock, K.H., Lancia, R.A., Conner, M.C., and Wood, B.L. 1985. A new change-in-ratio procedure robust to unequal catchability of types of animal. Biometrics, **41**: 653–662.
Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191.
Roseberry, J.L., and Woolfe, A. 1991. A comparative evaluation of techniques for analyzing white-tailed deer harvest. Wildl. Monogr. **117**: 1–59.
Routledge, R. 1989. The removal method for estimating natural populations: incorporating auxiliary information. Biometrics, **45**: 111–121.
Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd edition. Macmillan Publishing Co., New York, N.Y.
Udevitz, M.S., and Pollock, K.H. 1991. Change-in-ratio estimates for populations with more than two subclasses. Biometrics, **47**: 1531–1546.
Udevitz, M.S., and Pollock, K.H. 1995. Using effort information with change-in-ratio data for population estimation. Biometrics, **51**: 471–481.
Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics, **12**: 163–169.

Appendix 1. *Variance formulae for estimates of exploitation rate and catchability coefficient*

Here, we present formulae for estimating the variance of estimates of exploitation rate and catchability coefficient obtained from index-removal and change-in-ratio methods. As in the main text, it is assumed that the estimates of removals by type are independent of the pre- and post-season survey results. The exploitation rate on x-type animals can be estimated from index-removal analysis as

$$\hat{u}_{\text{xir}} = \frac{\hat{c}_{x1} - \hat{c}_{x2}}{\hat{c}_{x1}} .$$

The variance of this can be estimated as

$$\hat{V}(\hat{u}_{\text{xir}}) = \frac{\hat{c}_{x2}^{\,2}}{\hat{c}_{x1}^{\,4}} \hat{V}(\hat{c}_{x1}) + \frac{1}{\hat{c}_{x1}^{\,2}} \hat{V}(\hat{c}_{x2}) - 2\frac{\hat{c}_{x2}^{\,2}}{\hat{c}_{x1}^{\,3}} \widehat{\text{Cov}}(\hat{c}_{x1}, \hat{c}_{x2}).$$

Exploitation rate can be estimated from change-in-ratio analysis as

$$\hat{u}_{\text{xcir}} = \frac{\hat{f}(\hat{P}_1 - \hat{P}_2)}{\hat{f}\hat{P}_1 - \hat{P}_1\hat{P}_2} .$$

The variance can be estimated as

$$\hat{V}(\hat{u}_{\text{xcir}}) = \frac{\hat{f}^2\hat{P}_2^{\,2}}{\hat{P}_1^{\,4}(\hat{f}-\hat{P}_2)^2}\hat{V}(\hat{P}_1) + \frac{(\hat{P}_1-\hat{f})^2\hat{f}^2}{\hat{P}_1^{\,2}(\hat{f}-\hat{P}_2)^4}\hat{V}(\hat{P}_2)$$

$$+ \frac{(\hat{P}_2-\hat{P}_1)^2\hat{P}_2^{\,2}}{\hat{P}_1^{\,2}(\hat{f}-\hat{P}_2)^4}\hat{V}(\hat{f}) + 2\frac{\hat{f}^2\hat{P}_2(\hat{P}_1-\hat{f})}{\hat{P}_1^{\,3}(\hat{f}-\hat{P}_2)^3}\widehat{\text{Cov}}(\hat{P}_1, \hat{P}_2) .$$

The catchability coefficient for x-type animals can be estimated from index-removal analysis as

$$\hat{q}_{\text{ir}} = \frac{\hat{c}_{x1}}{\hat{X}_{\text{ir}}} = \frac{\hat{c}_{x1} - \hat{c}_{x2}}{\hat{R}_x}$$

and the variance can be estimated as

$$\hat{V}(\hat{q}_{\text{ir}}) = \frac{1}{\hat{R}_x^{\,2}}(\hat{V}(\hat{c}_{x1}) + \hat{V}(\hat{c}_{x2})) .$$

$$+ \frac{(\hat{c}_{x1}-\hat{c}_{x2})^2}{\hat{R}_x^{\,4}}\hat{V}(\hat{R}_x) - \frac{2}{\hat{R}_x^{\,2}}\widehat{\text{Cov}}(\hat{c}_{x1}, \hat{c}_{x2}) .$$

Finally, the catchability coefficient for x-type animals is obtained from change-in-ratio analysis as

$$\hat{q}_{\text{cir}} = \frac{\hat{c}_{x1}}{\hat{X}_{\text{cir}}} = \frac{\hat{c}_{x1}(\hat{P}_1 - \hat{P}_2)}{\hat{P}_1\hat{R}(\hat{f} - \hat{P}_2)}$$

with variance estimated as

$$\hat{V}(\hat{q}_{\text{cir}}) = \frac{\hat{q}_{\text{cir}}^{\,2}\hat{V}(\hat{c}_{x1})}{\hat{c}_{x1}^{\,2}} + \frac{\hat{q}_{\text{cir}}^{\,2}\hat{P}_2^{\,2}\hat{V}(\hat{P}_1)}{\hat{P}_1^{\,2}(\hat{P}_1-\hat{P}_2)^2} + \frac{\hat{c}_{x1}^{\,2}(\hat{P}_1-\hat{f})^2\hat{V}(\hat{P}_2)}{\hat{P}_1^{\,2}\hat{R}^2(\hat{f}-\hat{P}_2)^2}$$

$$+ \frac{\hat{q}_{\text{cir}}^{\,2}\hat{V}(\hat{R})}{\hat{R}^2} + \frac{\hat{q}_{\text{cir}}^{\,2}\hat{V}(\hat{f})}{(\hat{f}-\hat{P}_2)^2} + \frac{2\hat{c}_{x1}^{\,2}\hat{P}_2(\hat{P}_1-\hat{f})\widehat{\text{Cov}}(\hat{P}_1, \hat{P}_2)}{\hat{P}_1^{\,3}\hat{R}^2(\hat{f}-\hat{P}_2)^3}$$

$$+ \frac{2\hat{q}_{\text{cir}}\hat{P}_2\widehat{\text{Cov}}(\hat{c}_{x1}, \hat{P}_1)}{\hat{P}_1^{\,2}\hat{R}(\hat{f}-\hat{P}_2)} + \frac{2\hat{q}_{\text{cir}}(\hat{P}_1-\hat{f})\widehat{\text{Cov}}(\hat{c}_{x1}, \hat{P}_2)}{\hat{P}_1\hat{R}(\hat{f}-\hat{P}_2)^2}$$

$$+ \frac{2\hat{q}_{\text{cir}}^{\,2}\widehat{\text{Cov}}(\hat{f}, \hat{R})}{\hat{R}(\hat{f}-\hat{P}_2)} .$$

Appendix 2. *Biases of estimators of population parameters*

By the Taylor's series approach, the asymptotic bias of a function of estimators, $\boldsymbol{Y}$, is

$$\text{Bias}(g(\boldsymbol{Y})) = \sum_{i=1}^{n}\frac{1}{2}\frac{\partial^2 g}{\partial y_i^2}V(y_i) + \sum_{i<j}^{n}\frac{\partial^2 g}{\partial y_i\,\partial y_j}\text{Cov}(y_i, y_j).$$

This relationship can be used to derive the following expressions.

$$\text{Bias}(\hat{X}_{\text{cir}}) = \frac{N_1P_2}{(P_1-P_2)^2}V(\hat{P}_1) + \frac{N_2P_1}{(P_1-P_2)^2}V(\hat{P}_2)$$

$$- \frac{N_1P_1 + N_2P_2}{(P_1-P_2)^2}\text{Cov}(\hat{P}_1, \hat{P}_2) + \frac{P_1}{P_1-P_2}\text{Cov}(\hat{R}, \hat{f})$$

$$\text{Bias}(\hat{X}_{\text{ir}}) = \frac{X_2}{(c_{x1}-c_{x2})^2}V(\hat{c}_{x1}) + \frac{X_1}{(c_{x1}-c_{x2})^2}V(\hat{c}_{x2})$$

$$- \frac{X_1+X_2}{(c_{x1}-c_{x2})^2}\text{Cov}(\hat{c}_{x1}, \hat{c}_{x2})$$

$$\mathrm{Bias}\,(\hat{N}_{\mathrm{cir}}) = \frac{N_1}{(P_1 - P_2)^2} V(\hat{P}_1) + \frac{N_2}{(P_1 - P_2)^2} V(\hat{P}_2)$$

$$- \frac{N_1 + N_2}{(P_1 - P_2)^2} \mathrm{Cov}(\hat{P}_1, \hat{P}_2) + \frac{1}{P_1 - P_2} \mathrm{Cov}(\hat{R}, \hat{f})$$

$$\mathrm{Bias}\,(\hat{N}_{\mathrm{ir}}) = \frac{X_2}{P_1(c_{x1} - c_{x2})^2} V(\hat{c}_{x1}) + \frac{X_1}{P_1(c_{x1} - c_{x2})^2} V(\hat{c}_{x2})$$

$$+ \frac{X_1}{P_1^{\,3}} V(\hat{P}_1) - \frac{X_1 + X_2}{P_1(c_{x1} - c_{x2})^2} \mathrm{Cov}(\hat{c}_{x1}, \hat{c}_{x2})$$

$$+ \frac{X_2}{P_1^{\,2}(c_{x1} - c_{x2})} \mathrm{Cov}(\hat{P}_1, \hat{c}_{x1}) - \frac{X_1}{P_1^{\,2}(c_{x1} - c_{x2})} \mathrm{Cov}(\hat{P}_1, \hat{c}_{x2})$$

where here c_{xi} is the expected value of the catch rate of x-type animals in survey i. Estimates of the biases can be obtained by substituting estimates for the parameters in the above expressions.