

Influence of errors in natural mortality estimates in cohort analysis

G. Mertz and R.A. Myers

Abstract: The accuracy of the estimation of cohort strength from catch data may be greatly degraded if a poor estimate of the natural mortality rate is entered into the calculation. A straightforward, exact formulation for the error in cohort reconstruction due to a misspecified natural mortality rate is presented. The special case of constant fishing mortality is particularly transparent, allowing the error to be segmented into easily interpreted terms. A change in the fishing mortality may result in a distinct hump in the transient behavior of the bias factor, rather than a simple monotonic adjustment. This implies a similar pattern in estimated cohort strength.

Résumé : La précision de l'estimation de l'effectif d'une cohorte à partir des données de capture peut être réduite considérablement si on fait intervenir dans les calculs une estimation inexacte du taux de mortalité naturelle. On présente une formule exacte et simple pour corriger l'erreur dans la reconstruction d'une cohorte attribuable à un taux de mortalité naturelle mal précisé. Le cas spécial de la mortalité par pêche constante est particulièrement transparent, permettant la segmentation de l'erreur en termes facilement interprétables. Un changement dans la mortalité par pêche peut se traduire par une bosse distincte dans le comportement transitoire du facteur d'erreur, plutôt qu'en un simple ajustement monotone. Cela implique un changement semblable dans l'effectif estimé de la cohorte.

[Traduit par la Rédaction]

Introduction

The reconstruction of numbers-at-age from catch-at-age data is now a standard tool in fisheries science (Hilborn and Walters 1992). The techniques used for this reconstruction, e.g., cohort or virtual population analysis, are generically known as sequential population analysis (SPA). It is well appreciated that the result can be sensitive to the estimate of the natural mortality rate (M) assumed in the calculation. In this paper we provide a compact and illuminating expression for the error arising in SPA from the misspecification of M .

An appreciable number of investigations of errors in SPA precede this one: Pope (1972), Ulltang (1977), Sims (1984), Saila et al. (1985), Sampson (1988), Hilden (1988), Lapointe et al. (1989) and Akksland (1994). (This list is not intended to be exhaustive.) For our purposes we wish to highlight two results from these studies. First, implied by Sims (1984), Sampson (1988), and Akksland (1994), misspecification of M creates particularly large biases for lightly fished stocks. Second, changes in fishing mortality in the presence of a misspecified M , implied by Hilden (1988) and Lapointe et al. (1989), lead to a time-varying bias, thus imposing spurious time trends in SPA estimates of cohort size.

In regard to the first point, we show that it is possible to exactly formulate the error in cohort reconstruction, arising from misspecification of M , avoiding the Taylor series type

approximations or simulations employed in earlier studies. The case in which a cohort is subjected to a constant fishing mortality proves to be particularly enlightening, in that one can clearly separate different error terms and interpret them.

In regard to the second point, we show that bias arising from misspecified M may, in the presence of a change in the fishing mortality, exhibit a lobe-shaped time trend, rather than the monotonic adjustment one might anticipate. In the analysis that follows we are concerned with the errors that are common to all cohorts in a SPA, not those that arise from the difficulties in estimating the fishing mortality in the final year.

Analysis

Mertz and Myers (1996) showed that the population equation can be expressed as a generalized cohort analysis and from this follows a generalized catch equation. The generalized cohort analysis is given by (a related equation may be found in Hiramatsu 1995):

$$(1) \quad N_{i+1} - e^{-M}N_i = -\gamma_i(M)C_i$$

where N_i is the number of fish in the cohort in year i , C_i is the catch in year i , M is the natural mortality rate, and

$$(2) \quad \gamma_i(M) = \frac{e^{-M} \int_0^1 e^{Mt} C'_i(t) dt}{C_i}$$

Here, $C'_i(t)$ represents the instantaneous catch per unit time in year i , $t = 0$ represents the beginning of any year, and $t = 1$ its end. (Note that $C_i = \int_0^1 C'_i(t) dt$.) Mertz and Myers (1996) showed that, corresponding to eq. 1, there is a generalized catch equation given by

$$(3) \quad C_i = \gamma_i(M)^{-1} N_i (e^{-M} - e^{-(F_i+M)})$$

where F_i is the fishing mortality for year i , which is implicitly defined by $N_{i+1} = N_i \exp[-(F_i + M)]$. If fishing mortality does

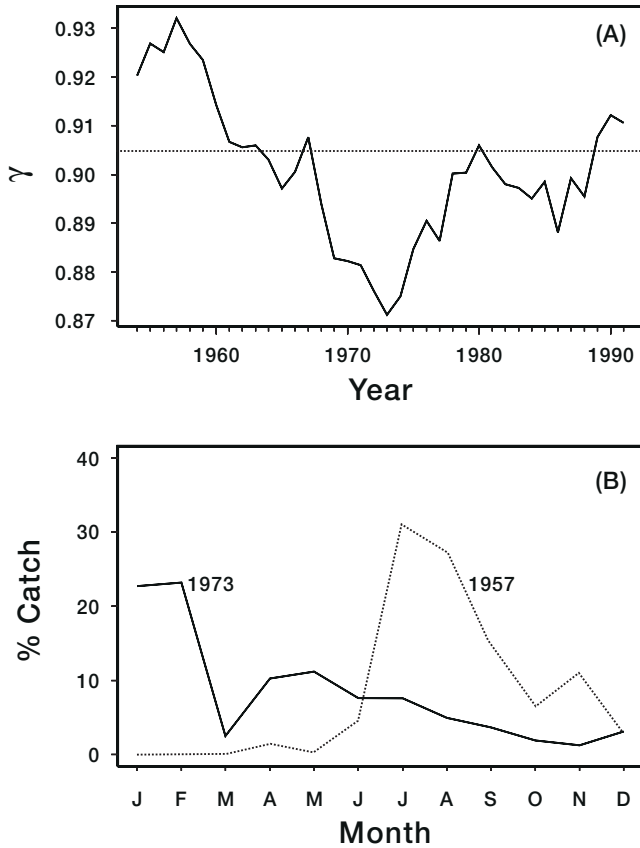
Received July 9, 1996. Accepted January 24, 1997.
J13542

G. Mertz.¹ Science Branch, Department of Fisheries and Oceans, Northwest Atlantic Fisheries Centre, P.O. Box 5667, St. John's, NF A1C 5X1, Canada.

R.A. Myers. Killiam Chair in Ocean Studies, Department of Biology, Dalhousie University, Halifax, NS B3H 4J1, Canada.

¹ Author to whom all correspondence should be addressed.
e-mail: mertz@athena.nwafc.nf.ca

Fig. 1. (A) The factor $\gamma_i(M)$ calculated from eq. 2 using monthly catch data for northern cod, i.e., cod off Labrador and the northern Grand Banks. The dotted line is Pope's approximation. (B) Monthly catches expressed as percentages of the total for the year at the maximum and minimum $\gamma_i(M)$. Note that this figure in Mertz and Myers (1996) had the labels "1957" and "1973" reversed.



not change within a year, the standard Baranov result can be recovered from eqs. 2 and 3 (in other words, Baranov is a special case of this generalized cohort analysis; Mertz and Myers 1996).

This formulation is preferable because any pattern of seasonal fishing can be treated as a special case. For our purposes, it leads to a simple formulation in which different errors can be understood without resorting to complicated approximations, iterative solutions of the Baranov equation, or simulations.

The quantity $\gamma_i(M)$ may be analytically calculated when C'_i has a simple seasonal dependence. For example, if one has a pulse fishery at time $t = \alpha$ ($0 \leq \alpha \leq 1$), then $\gamma_i(M) = \gamma(M) = (e^{-M(1-\alpha)})$. For $\alpha = 0.5$, one recovers the classical cohort analysis of Pope (1972). More generally, $\gamma_i(M)$ may be computed numerically, from eq. 2, using monthly catch-at-age data. Here we are tacitly assuming that the seasonal pattern of catches for any given age is adequately represented by the seasonal pattern of the age-aggregated catch, an assumption that will generally be very accurate (Mertz and Myers 1996, p. 161). If M is not large the γ_i values show great interannual stability. This is indicated in Fig. 1, the case of northern (Northwest Atlantic Fisheries Organization (NAFO) area 2J3KL) Atlantic cod

(*Gadus morhua*) for which the $\gamma_i(M = 0.2)$ series shows very little fluctuation despite the marked changes in seasonal catch pattern.

If the initial size of a cohort is N_1 , it may be expressed in terms of the catches using the following relation, derived from eq. 1:

$$(4) \quad N_1 = \gamma_1(M)e^M C_1 + \gamma_2(M)e^{2M} C_2 + \dots + \gamma_i(M)e^{iM} C_i + \dots$$

Thus, N_1 is easily recovered from the catch data when M is known. In practice, the initial cohort size is computed using an estimate of M , which we designate as \hat{M} . Now, the estimate of N_1 , designated as \hat{N}_1 , is given by

$$(5) \quad \hat{N}_1 = \gamma_1(\hat{M})e^{\hat{M}} C_1 + \gamma_2(\hat{M})e^{2\hat{M}} C_2 + \dots + \gamma_i(\hat{M})e^{i\hat{M}} C_i + \dots$$

Constant fishing mortality

Before deriving a general expression for the error in N_1 associated with an incorrect M , we will set the stage by examining the case of constant fishing mortality ($F_i = F$), which yields some helpful insights. In practical terms, a constant F is meaningful in the context of classical catch-curve analysis (fitting a line to logcatch vs. time, e.g., Hilborn and Walters 1992). In this case, we have $C_2 = C_1 e^{-(F+M)}$, $C_3 = C_1 e^{-2(F+M)}$, etc. Substituting these expressions into eq. 5 yields

$$(6) \quad \hat{N}_1 = \gamma(\hat{M})e^{\hat{M}} C_1 (1 + e^{(\hat{M}-M)-F} + \dots + e^{i[(\hat{M}-M)-F]} + \dots)$$

It follows directly from eq. 3 that $C_1 = \gamma(M)^{-1} e^{-M} N_1 (1 - e^{-F})$. We can use this expression to eliminate C_1 in eq. 6 and then sum the geometric series. The result is

$$(7) \quad \frac{\hat{N}_1}{N_1} = \frac{\gamma(\hat{M})e^{\hat{M}}}{\gamma(M)e^M} R(\Delta M, F)$$

where $\Delta M \equiv \hat{M} - M$, and

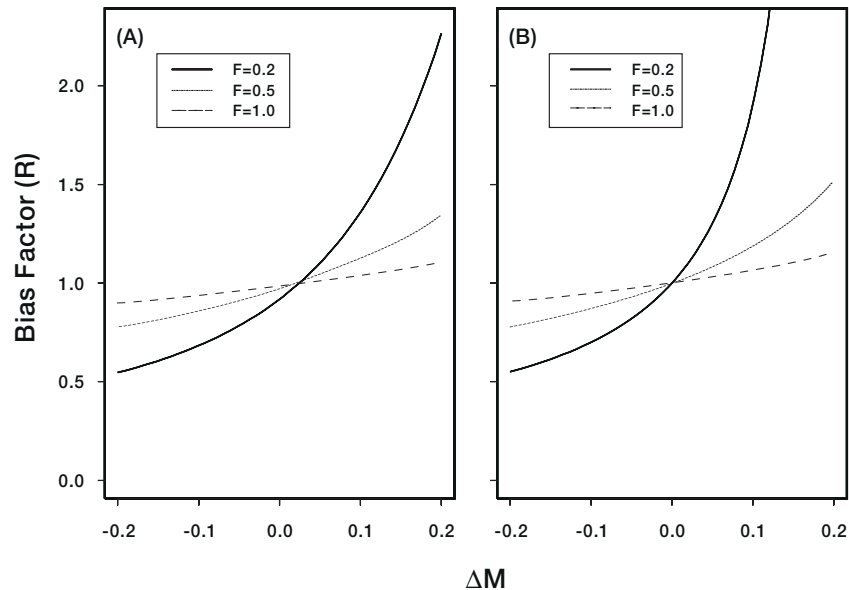
$$(8) \quad R(\Delta M, F) \equiv \frac{(1 - e^{-F})(1 - e^{n(\Delta M - F)})}{1 - e^{\Delta M - F}}$$

where n is the number of terms summed. The reasons for expressing eq. 7 in this form will become apparent.

The presence, in eq. 7, of the term $\gamma(\hat{M})e^{\hat{M}}/(\gamma(M)e^M)$ can be understood from the following example. In the limit of infinite fishing mortality, $R(\Delta M, F) = 1$ and all of the discrepancy between \hat{N}_1 and N_1 can be attributed to the term $\gamma(\hat{M})e^{\hat{M}}/(\gamma(M)e^M)$. Consider a pulse-type fishery (still with infinite fishing mortality) at time α (again $0 \leq \alpha \leq 1$). At the time of the catch, we know the exact cohort size is equal to C_1 . At the beginning of the year, then, the cohort size was $e^{M\alpha} C_1$. For a pulse fishery (as noted earlier), $\gamma(M) = e^{-(1-\alpha)M}$, so that $e^M \gamma(M) = e^{M\alpha}$, and thus it is clear that the term $\gamma(M)e^M$ in eq. 7 arises from the projection of the cohort size at the time of catch back to the beginning of the year. Thus, the term $\gamma(\hat{M})e^{\hat{M}}/(\gamma(M)e^M)$ represents the uncertainty arising from making this projection with a misspecified M . In fact this source of error is quite innocuous in that the maximum value of this term (attained when the entire fishery occurs as a pulse at year end) is given by $e^{\Delta M}$. Thus, the associated fractional error in N_1 , $\hat{N}_1/N_1 - 1$, cannot be larger than $e^{\Delta M} - 1 \approx \Delta M$. (The corresponding expression for a pulse fishery at time α is $\hat{N}_1/N_1 - 1 = e^{\alpha\Delta M} - 1 \approx \alpha\Delta M$.)

Before discussing the effect of ΔM on $R(\Delta M, F)$, it must be noted that $\Delta M = 0$ does not guarantee $R(\Delta M, F) = 1$; in fact $R(0, F) = 1 - e^{-nF}$. The departure from unity represents a

Fig. 2. The cumulative error in the ratio of the estimated to the true cohort size, $R(\Delta M, F)$, as a function of the difference in the estimated and the natural mortality, $\Delta M \equiv \hat{M} - M$, and a constant fishing mortality, F (eq. 8). The total number of catches-at-age entered to estimate the cohort size (n) is either $n = 5/(M + F)$ (A) or $n = \infty$ (B). The difference between the right and left panels is caused by different truncation effects associated with the number of ages used in the calculation.



truncation error, arising from the fact that only a finite number of annual catches are entered in the cohort analysis. This problem arises in real situations, in that reliable samples from the catches will be available only up to a given age, necessitating the assumption that there were no fish extant beyond that age (e.g., Hilborn and Walters 1992). An alternative procedure is to estimate the fishing mortality on the oldest fish and correct, using $R(0, F)$, for the truncation error (e.g., Myers and Cadigan 1995). The analysis above shows that this is only possible if ΔM is negligible; in general, one cannot separate the truncation error from the influence of ΔM (because ΔM is unknown).

The term $R(\Delta M, F)$ can inflate rapidly with increases in ΔM . Note that if $\Delta M \geq F$, then $R(\Delta M, F)$ diverges as n becomes very large (if $\Delta M < F$ then $R(\Delta M, F)$ converges to a finite value as n grows). The reason for this is clear from eq. 6: when $\Delta M \geq F$, each term in eq. 6 is greater than or equal to 1, so that the sum must diverge. A simple special case is $\Delta M = F$ for which it follows that $R(\Delta M, F) = n(1 - e^{-F}) \approx nF$; this approximation holds when F is small. Thus, if $n = 10$ and $F = 0.2$, $R(\Delta M, F) \approx 2$, which is twice the ideal. It is apparent that the most serious errors in cohort reconstruction will arise from this term.

In Fig. 2 we display $R(\Delta M, F)$ as a function of ΔM for three F values. In Fig. 2A we have specified $n = 5/(F + M)$ (so that the cohort has decayed by a factor of $e^{-5} = 0.007$) and $M = 0.2$. For comparison, Fig. 2B displays the same examples, but with $n = \infty$, for which there is no truncation effect and which is completely independent of M . It is evident that for $F = 0.5$ and 1.0 the departures of $R(\Delta M, F)$ from 1 are tolerable. For $F = 0.2$, misspecification of M can lead to severe errors in the cohort reconstruction, particularly if M is overestimated (a result foreshadowed above). Divergent behavior of $R(\Delta M, F)$ for $F = 0.2$, as ΔM approaches F , is evident in Fig. 2B. This is a

manifestation of the strong dependence of $R(\Delta M, F)$ on n when $\Delta M = F$, noted above. In other words, when ΔM becomes comparable with F , the result of the cohort analysis will be very sensitive to the number of catches-at-age summed.

A general expression

The results presented above hold qualitatively when the fishing mortality is not fixed over the lifetime of a cohort. To demonstrate this we utilize the following relation, derived from eqs. 1 and 3:

$$(9) \quad C_i = \gamma_i(M)^{-1} N_1 e^{-iM} e^{-(F_1 + F_2 + \dots + F_{i-1})} (1 - e^{-F_i})$$

We now enter the catches as specified by eq. 9 into eq. 5 and find

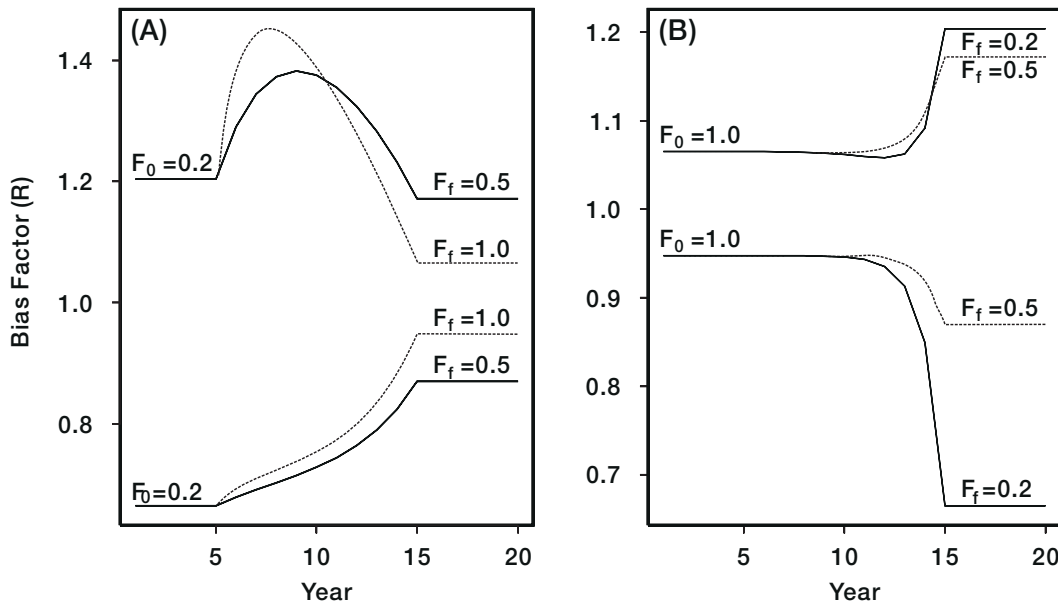
$$(10) \quad \hat{N}_1 = N_1 \sum_{i=1}^{i=n} \frac{\gamma_i(\hat{M})}{\gamma_i(M)} e^{\Delta M i} e^{-(F_1 + F_2 + \dots + F_{i-1})} (1 - e^{-F_i})$$

If ΔM is comparable with the typical fishing mortality experienced by the cohort, the terms within the summation will not decrease in magnitude as i increases, leading to gross inflation of the estimate of the cohort size. This result is a close parallel to that found for the constant fishing mortality case.

Spurious time trends

The bias created by a misspecified M is dependent on the fishing mortality, and it follows that a changing F will cause a time-varying bias, which will create spurious trends in the estimates of cohort strength. Lapointe et al. (1989) have demonstrated this effect. Using the tools above we can analytically

Fig. 3. (A) The effect of increasing fishing mortality ($F_0 = 0.2$) to $F_f = 0.5$ or $F_f = 1.0$ in year 5 on the ratio of the estimated to the true recruitment as manifested through $R(\Delta M, F_0, F_f)$. The top two lines correspond to $\Delta M = 0.1$, whereas the bottom two lines correspond to $\Delta M = -0.1$. Ten ages are used in the population reconstruction. (B) Same as (A) except that fishing mortality is decreased from $F_0 = 1.0$ to $F_f = 0.2$ or $F_f = 0.5$.



solve the case where a cohort is fished for j years at F_0 and for $n - j$ years at the final fishing mortality (F_f) (where n is again the range of ages entered into the cohort reconstruction). We assume that $\gamma_i(M)$ does not change substantially when F changes, an approximation that is often very accurate (Mertz and Myers 1996). Thus, the influence of the changing fishing mortality is manifested through the R term ($R(\Delta M, F_0, F_f)$, formally). Using either eq. 5 or eq. 10, we find

$$(11) \quad R(\Delta M, F_0, F_f) = \frac{(1 - e^{-F_0})(1 - e^{(\Delta M - F_0)j})}{1 - e^{(\Delta M - F_0)}} + \frac{(1 - e^{-F_f})e^{(\Delta M - F_0)j}(1 - e^{(\Delta M - F_f)(n-j)})}{1 - e^{(\Delta M - F_f)}}$$

Below, we use eq. 11 to demonstrate the effect of changing fishing mortality on cohort reconstruction.

An application of eq. 11 is shown in Fig. 3. We have considered a species for which 10 different ages are present in the fishery ($n = 10$). In each year a new cohort becomes finished, in the terminology of Evans (1996), that is, it has been fished for 10 years. The plotted R values refer to these newly finished cohorts. The fishing mortality is constant at F_0 until and including year 5, with a transition to F_f in year 6, with no further changes in F . In year 5 and all prior years, the newly finished cohorts have all been fished at F_0 , i.e., $j = 10$ in eq. 11. In year 6, the newly finished cohort has been fished at F_0 for 9 years ($j = 9$), and at F_f for 1 year ($n - j = 1$). In year 15 and beyond, $j = 0$, and each newly finished cohort has been fished entirely at F_f .

In the examples plotted in Fig. 3, R exhibits a time trend as it adjusts from the first level of fishing to the second. Since R gives the bias in the cohort reconstruction, it is evident that the bias will vary temporally, which will introduce spurious trends into the estimates of cohort strength. (Strictly, R represents the

most important source of bias, rather than the entire bias.) Particularly interesting is the case where ΔM is positive and fishing mortality increases; the bias actually increases before declining to the value appropriate to F_f . Thus, the trend induced by changing bias need not be a simple monotonic adjustment to the level set by the final fishing mortality. Note that an analogous result was obtained by Hilden (1988) for cases in which unfinished cohorts are reconstructed with a tuned virtual population analysis. Our analysis has shown that these effects exist even when a tuning procedure is not applied (that is, when only finished cohorts are dealt with).

Discussion

We have dealt with one aspect of the errors arising in SPA, that due to misspecification of M . The matter of estimating the size of a cohort prior to its complete removal is not dealt with here. This problem requires statistical fitting of the available catch-at-age data to a suitable index, which is itself a potential source of large errors (Hilborn and Walters 1992; Myers and Cadigan 1995). We have also neglected errors in catch (e.g., Rivard 1989) and aging errors (e.g., Bradford 1991). There are, however, clear advantages to isolating the effect of misspecification of M .

Our analysis has yielded an exact, general equation for the bias incurred by misspecification of M in cohort reconstruction (eq. 10). The case of constant fishing mortality has proved to be particularly enlightening, in that the error can be segregated into easily interpretable terms, these being the truncation error, the projection error, and the cumulative error (that resulting from the sum over ages). Finally, we showed that there are unanticipated aspects of the trends in bias arising from a changing fishing mortality in confluence with a misspecified M . We briefly discuss the impact of these error influences below.

The truncation error will be harmless for most fish stocks. It is potentially correctable but only if one has confidence that ΔM is small so that $R(0, F)$ can be used to make the correction.

The projection error results from transposing the population estimate from the time of the catch to the beginning of the year. As noted above this error, in fractional terms, is given by $\hat{N}_1/N_1 - 1 = e^{\alpha\Delta M} - 1$ for a pulse catch at time α (which lies in the interval 0–1). Because Pope's (1972) cohort analysis is often adequate for population reconstruction and because it is equivalent to $\alpha = 0.5$, the projection error will typically be $e^{0.5\Delta M} - 1 \approx 0.5\Delta M$. Thus, it is unlikely that the projection error will exceed 10%.

The summation over ages in the cohort reconstruction can lead to a serious cumulative error when there is an overestimate in M which is comparable with the fishing mortality. In fractional terms, the severity of the error when M is underestimated is not as great, but it can nevertheless be appreciable. Based on our analysis (Fig. 2 in particular) we would suggest that $|\Delta M| \leq 0.5F$ is necessary for accurate cohort reconstruction. This implies that stocks that experience low fishing mortalities, in the range of 0.1–0.2, are particularly vulnerable to gross errors of estimation of the cohort size.

If fishing mortality is constant, the errors referred to above may not be very detrimental, in that the estimate of cohort size will simply be distorted by a constant factor. (We are assuming that natural mortality is constant.) If one is analyzing time series of log-transformed recruitment from cohort analysis, the error will simply appear as a constant offset and will not influence trends or estimates of the standard deviation of the log-transformed recruitment, a useful index of recruitment variability.

If fishing mortality does vary appreciably, then the situation is more complicated. The bias engendered by misspecification of M will vary temporally, introducing deceptive trends into the reconstructed series of recruitments. The artifactual trend may contain a hump, rather than being a monotonic adjustment from one level of bias to another, a somewhat surprising result.

Acknowledgments

We thank Nick Barrowman and Noel Cadigan for comments on the manuscript.

References

- Aksland, M. 1994. A general cohort analysis method. *Biometrics*, **50**: 917–932.
- Bradford, M.J. 1991. Effects of ageing errors on recruitment time series estimated from sequential population analysis. *Can. J. Fish. Aquat. Sci.* **48**: 555–558.
- Evans, G.T. 1996. Using elementary operations of sequential population analysis to display problems in catch or survey data. *Can. J. Fish. Aquat. Sci.* **53**: 239–243.
- Hiramatsu, K. 1995. A theoretical study of equations used in virtual population analysis. *Fish. Sci.* **61**: 752–754.
- Hilborn, R., and Walters, C.J. 1992. *Quantitative fisheries stock assessment: choice, dynamics and uncertainty*. Chapman & Hall, New York.
- Hilden, M. 1988. Errors in perception in stock and recruitment studies due to wrong choices of natural mortality rate in virtual population analysis. *J. Cons. Int. Explor. Mer.* **44**: 123–134.
- Lapointe, M.F., Peterman, R.M., and MacCall, A.D. 1989. Trends in fishing mortality rate along with errors in natural mortality rate can cause spurious time trends in fish stock abundances estimated by virtual population analysis (VPA). *Can. J. Fish. Aquat. Sci.* **46**: 2129–2139.
- Mertz, G., and Myers, R.A. 1996. An extended cohort analysis: incorporating the effect of seasonal catches. *Can. J. Fish. Aquat. Sci.* **53**: 159–163.
- Myers, R.A., and Cadigan, N.G. 1995. Statistical analysis of catch-at-age data with correlated errors. *Can. J. Fish. Aquat. Sci.* **52**: 1265–1273.
- Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. *Int. Comm. Northwest Atl. Fish. Res. Bull.* **9**: 65–74.
- Rivard, D. 1989. Overview of the systematic, structural and sampling errors in cohort analysis. *Am. Fish. Soc. Symp.* **6**: 49–65.
- Saila, S.B., Lorda, E., and Walker, H.A. 1985. The analysis of parameter error propagation in simple fishery models. *Mar. Res. Econ.* **1**: 235–246.
- Sampson, D.B. 1988. The stability of virtual population analysis cohort size estimates. *J. Cons. Int. Explor. Mer.* **44**: 135–142.
- Sims, S.E. 1984. An analysis of the effect of errors in the natural mortality rate on stock-size estimates using virtual population analysis (cohort analysis). *J. Cons. Int. Explor. Mer.* **41**: 149–153.
- Ulltang, O. 1977. Sources of errors in and limitations of virtual population analysis (cohort analysis). *J. Cons. Int. Explor. Mer.* **37**: 249–260.