Abstract.–Determining absolute survival rates for larval fishes is extremely difficult. However, many ecological questions concern relative survival of two groups. For example, we might ask: (1) Do older larvae have higher survival than younger larvae? and (2) Do faster growers have higher survival than slower growers? We present a simple model and several estimation schemes for the ratio of survival rates based on monitoring relative abundance of the two groups over time. When the logarithm of the ratio of abundances is regressed on time, the resulting estimate of slope is an estimate of the difference in instantaneous mortality rates. An estimate of the ratio of survival rates is obtained by exponentiating the slope. The model is shown to be a logistic model and can be fitted by maximum likelihood methods.

Estimating Relative Survival Rate for Two Groups of Larval Fishes from Field Data: Do Older Larvae Survive Better Than Young?

John M. Hoenig
Pierre Pepin

Science Branch, Department of Fisheries and Oceans
P.O. Box 5667, St. John's, Newfoundland, A1C 5X1, Canada

William D. Lawing

Departments of Statistics and Industrial Engineering
University of Rhode Island, Kingston, Rhode Island 02881

Determining absolute survival rates for larval fishes is extremely difficult, even when cohorts can be accurately identified by means of daily growth rings in the otoliths. For example, survival between times 1 and 2 might be estimated by catch/tow at time 2 ÷ catch/tow at time 1. Due to random sampling error alone, this estimate may be nonsensical (>1), and the chances of obtaining nonsensical estimates increases with increasing patchiness in the distribution of larvae over space.

In some cases, it may not be necessary to estimate absolute survival rates. Estimates of relative survival rate may be sufficient and easier to obtain. Many ecological questions concern the relative survival rates of two or more groups. For example, at a given point in time, the older larvae present should be larger than the younger larvae and, hence, may have a higher survival rate (Peterson and Wroblewski 1984, McGurk 1986). On the other hand, it has been suggested that larvae born later (i.e., younger larvae) may have a higher survival than larvae born earlier (Victor 1983, Methot 1983, Crecco and Savoy 1985, Rice et al. 1987). Also, it has been suggested that faster-growing larvae survive better than slower-growing ones (Rosenberg and Haugen 1982).

In this paper, we consider how relative survival rates can be estimated for two groups occurring at the same time and place from field data consisting of the composition of the catch at two or more times. The intuitive basis for the method is this: Changes in the relative abundance of two groups over time reflect differences in mortality rates (assuming no emigration or immigration occurs). The methods we present allow for catchability to differ for the two groups and to vary over time. However, the relative catchability (ratio of the catchability coefficients) of the groups cannot change over the time period considered. Thus, if factors such as wind, currents, boat speed, or net clogging vary among sampling periods, then the same proportional change in the catchability coefficient must occur for both groups. This is less restrictive than assuming constant catchability over time (as when estimating absolute survival rate by the decline in catch-per-unit-sampling-effort over time).

Development of the model

Suppose that at time t, a sample of larvae is obtained and that examination of the otoliths reveals that all larvae are of the same (approximate)

age. It is also determined that larvae can be classified as either fast or slow growing on the basis of the widths of the growth increments. Thus, we can determine the proportion of the larvae that are fast growing. If we sample the same population at time $t + 1$, we can again separate the larvae into fast- and slow-growing groups on the basis of the width of the otolith at time t (not at time $t + 1$).

Alternatively, suppose that a sample of larvae at time t has two cohorts, where a cohort is defined to be all larvae hatching in a given week. We can follow the relative abundance of the two cohorts in the catch over time by taking repeated samples from the population.

We will let the subscripts E and L refer to the two cohorts (e.g., E for early- and L for late-spawned). Let the number of larvae in the population's two cohorts at time t be N_{Et} and N_{Lt}. Suppose that the size of each cohort declines exponentially over time such that, for the ith cohort ($i \in \{E,L\}$),

$$N_{it} = N_{i0} \exp(-Z_i t)$$

where Z_i is the instantaneous mortality rate (time^{-1}) and N_{i0} is the initial abundance of the ith group. Also suppose that the expected catch of animals from group i at time t (C_{it}), for a standard unit of effort, is proportional to the abundance, i.e.,

$$C_{it} = q_{it} N_{it}$$

where q_{it} is the time- and group-specific catchability coefficient. Then the ratio of expected catches, R_t, is

$$R_t = \frac{C_{Lt}}{C_{Et}} = \frac{q_{Lt} N_{L0} \exp(-Z_L t)}{q_{Et} N_{E0} \exp(-Z_E t)}. \tag{1}$$

We assume that the ratio of the catchability coefficients (q_{Lt}/q_{Et}) is constant over the course of the study. Since the ratio of initial abundances (N_{L0}/N_{E0}) is also a constant, equation (1) can be rewritten

$$R_t = q \frac{\exp(-Z_L t)}{\exp(-Z_E t)} = q \exp\{(Z_E - Z_L)t\} \tag{2}$$

where q is a nuisance parameter that subsumes the catchability coefficients and initial abundances. Taking logarithms of (2) results in a linear relationship with respect to time:

$$\log_e(R_t) = \log_e(q) + (Z_E - Z_L)t. \tag{3}$$

Thus, regressing the logarithm of the observed ratio of abundances ($\hat{R}_t$) against time results in a linear relationship with slope equal to an estimate of the dif-

ference in the instantaneous mortality rates. (The proper weighting to use in a weighted regression is discussed below.) Note that it is sometimes necessary to add a small constant to the numerator and denominator to avoid dividing by 0 or taking the logarithm of 0.

Exponentiating the slope estimated by (3) provides an estimate of the ratio of the finite survival rates:

$$\widehat{S_L/S_E} = e^{slope}$$

where $S_L = e^{-Z_L}$; $S_E = e^{-Z_E}$.

By the Taylor's series (delta) method, the asymptotic variance of the estimated survival ratio can be approximated by (Seber 1982):

$$V(\widehat{S_L/S_E}) = e^{(2\ slope)} V(slope).$$

Diagnostics

Under the assumptions given above (constant ratio of catchabilities over time and constant difference in instantaneous mortality rates), a plot of the logarithm of the ratio of catches versus time would be expected to be linear (assuming the sample sizes are reasonably large). A departure from linearity suggests violation of one or both of the assumptions. This provides a diagnostic procedure to check on the assumptions. If catchabilities vary in a nonsystematic fashion, the fit of the regressions would be low and would not be influenced by increasing sample size.

Two-sample estimator

A special case is when only two samples have been obtained. Then, exponentiating the slope of the line described by equation (3) reduces to the change-in-ratio estimator of relative survival described by Paulik and Robson (1969). Thus,

$$\left(\widehat{\frac{S_L}{S_E}}\right) = \frac{\hat{R}_2}{\hat{R}_1} = \frac{\hat{C}_{L2}\ \hat{C}_{E1}}{\hat{C}_{L1}\ \hat{C}_{E2}} \tag{4}$$

where the $\frown$ symbol indicates estimated quantities. If the proportion of early spawners in samples 1 and 2 are denoted by P_1 and P_2, respectively, then an estimate of the variance can be found by the Taylor's series method to be (Seber 1982, p. 382)

$$V(\widehat{S_L/S_E}) = [P_1(1 - P_2)]^{-4}$$

$$\{(1 - P_2)^2\ P_2{}^2\ V(P_1)$$

$$+ (1 - P_1)^2\ P_1{}^2\ V(P_2)\}.$$

Here, $V(P_1)$ and $V(P_2)$ are the variances of the proportions, found by the usual formula for variance of a binomial, i.e.,

$$V(P_i) = \frac{P_i(1-P_i)}{N_i}$$

where N_i is the number of larvae in sample i, and i can take on the values of 1 and 2.

Weighted regression

Equation (3) can be seen to be a logistic model by noting that $\log_e(R_t)$ is the logistic transformation. That is, letting P_t be the proportion in the sample at time t which belongs to group "L" (i.e., $\frac{C_L}{C_L+C_E}$), the logit transformation of the proportion is

$$\text{logit}(P_t) = \log_e\left(\frac{P_t}{1-P_t}\right) = \log_e(R_t).$$

This allows us to refer to known results for logistic models to determine the optimum weighting scheme.

Not all observations on catch composition are of equal value in estimating the relative survival rate. This is because the variance of the ratio of catches for any particular sampling date will be a function of the sample size and the proportions in the population. One method for specifying weights to be used in the regression, which explicitly accounts for this, is:

$$\text{weight}_t = \hat{C}_{Et}^{-1} + \hat{C}_{Lt}^{-1}.$$

When these weights are used in the regression, the resulting estimates are known as minimum logit chi-square estimates. Once the logistic regression model has been fitted, new weights can be computed using the predicted catch composition, rather than the observed composition. The regression can then be recomputed, the weights updated, and the regression recomputed. until adequate convergence is achieved. This procedure results in maximum likelihood estimates (McCullagh and Nelder 1983). Most standard statistical packages will perform logistic regression so that the user need not specify explicitly the iterative weighting scheme.

The above weighting scheme is appropriate when larvae are sampled randomly, i.e., each larva is sampled independently of all other larvae. This situation is approximated when the expected catch per tow is small (≤ 1) and tows are random over space. When the expected catches are large, the sampling procedure results in cluster samples. Consequently, the theoretical binomial variance is too small. However, the binomial variance becomes increasingly small as the sample size increases and as the proportions approach the extremes (0 or 1), and this is what would be expected for cluster sampling. Thus, weights computed from the binomial variance should be reasonably appropriate.

Estimation when there are many small samples

It sometimes occurs that catch rates are extremely low and many small samples are obtained. For example, sampling may be conducted daily with low intensity. In this case, the logistic regression procedure will not work well due to the occurrence of many zero catches. Estimation under the logistic model can still be accomplished by constructing the likelihood function and solving directly for the difference in mortality rates that maximizes the likelihood function.

The probability that an individual, randomly selected from the catch at time t, is from group E (early-spawning group) is equal to

$$\begin{aligned} P_t(E) &= \frac{\hat{C}_{Et}}{\hat{C}_{Et}+\hat{C}_{Lt}} \\[2mm] &= \frac{q_{Et}\,N_{E0}\,e^{-Z_E t}}{q_{Et}\,N_{E0}\,e^{-Z_E t} + q_{Lt}\,N_{L0}\,e^{-Z_L t}} \\[2mm] &= \frac{1}{1 + qe^{\Delta Z t}} \end{aligned} \tag{5}$$

where q is the "apparent" initial ratio of abundances (actual initial ratio of abundances if the ratio of catchability coefficients is equal to 1) and ΔZ is the difference in instantaneous mortality rates ($Z_E - Z_L$).

The probability that an animal is from group L is the complement of (5):

$$\begin{aligned} P_t(L) &= 1 - \frac{1}{1 + qe^{\Delta Z t}} \\[2mm] &= \frac{qe^{\Delta Z t}}{1 + qe^{\Delta Z t}}. \end{aligned}$$

The likelihood function can then be constructed as the product of the probabilities for each animal in each sample:

$$\begin{aligned} \Lambda &= \prod_{i=1}^{n_E} \frac{1}{1 + qe^{\Delta Z t_i}} \prod_{j=1}^{n_L} \frac{qe^{\Delta Z t_j}}{1 + qe^{\Delta Z t_j}} \\[2mm] &= \frac{q^{n_L}\,e^{(\Delta Z \sum_{j=1}^{n_L} t_j)}}{\prod_{i=1}^{n_E+n_L} (1 + qe^{\Delta Z t_i})} \end{aligned} \tag{6}$$
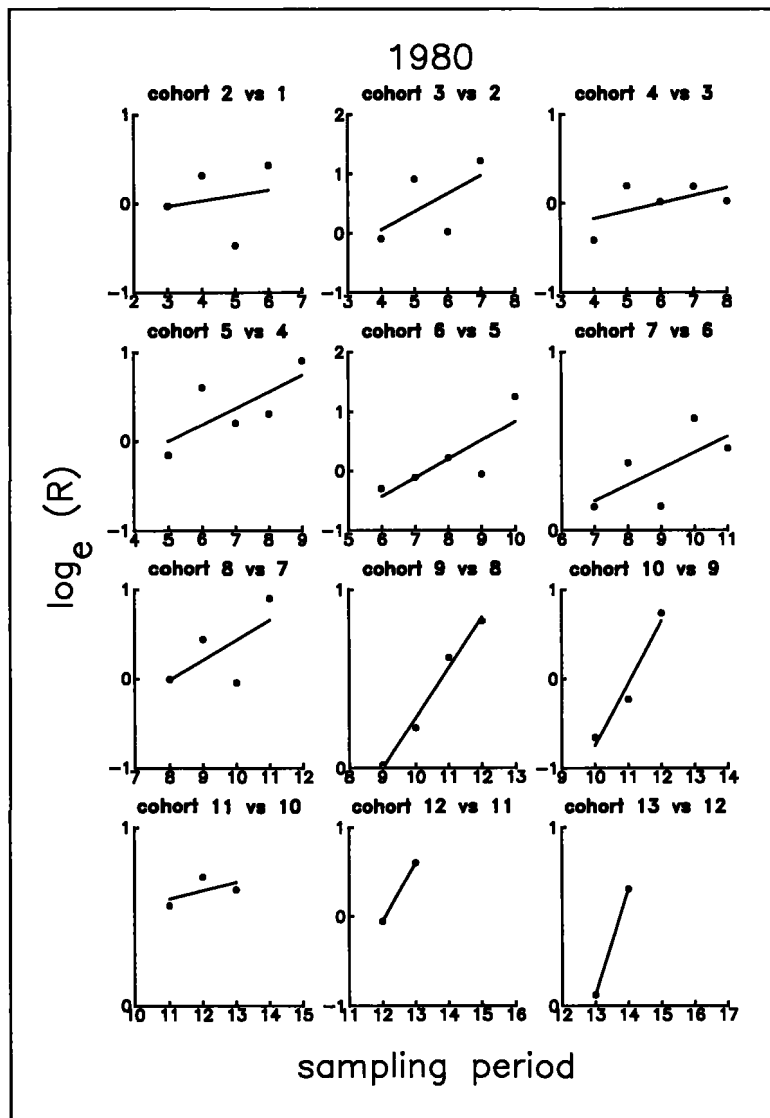
Figure 1

Plots of the logarithm of the ratio of abundance (cohort $i+1$ ÷ cohort i) versus sampling period (measured as 5-day intervals) in 1980. Slopes of the linear regression lines estimate $Z_E - Z_L$. (Note that the scale of the ordinate varies among plots.) In all 12 comparisons of cohorts, the slope is positive (later cohort has the better survival).

where n_E and n_L are the total number of animals observed in groups E and L, respectively, t_i is the time when the ith individual from group 1 was caught, and similarly for t_j. The major statistical packages have routines for maximizing this function.

Generalizations

The logistic model is easily extended to enable one to incorporate the effects of covariates. For example, suppose one has information on the time elapsed since the sampling program began (t) and also on the number of days (t *) a power plant, which can be an additional source of mortality, has been operating for each of the sampling dates. Then instead of having t as the only explanatory variable, we can have a linear combination of explanatory variables in the exponent

$$\Delta Z t_i + b_1 \ t_i^*$$

where b_1 is the differential mortality rate (per time) of the two groups attributable to power plant operation.

Example

We consider data on American shad *Alosa sapidissima* in the Connecticut River, Connecticut, USA, kindly supplied by Victor Crecco and Thomas Savoy (Dep. Environ. Prot., Mar. Fish Office, Waterford, CT). A description of the study area and sampling methods, and a careful analysis of the data, are given in Crecco et al. (1983). Our purpose in considering these data is to illustrate the use of our method and to explore the kinds of questions that can be asked by study-
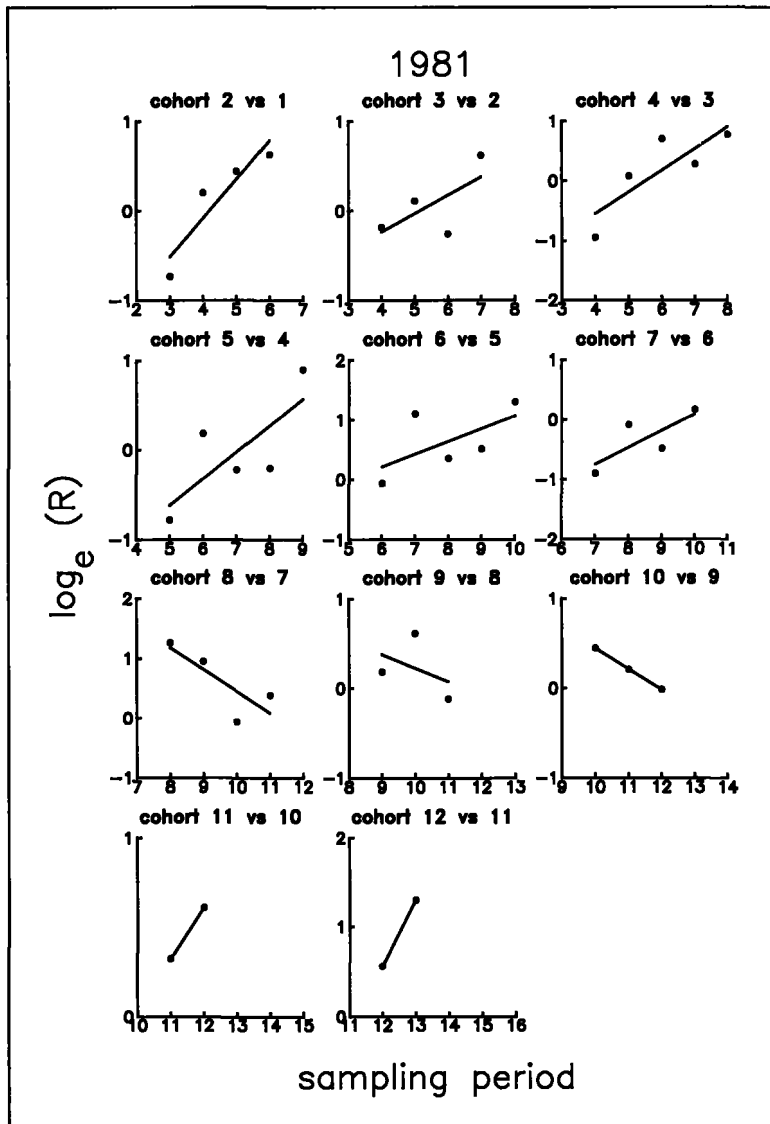
Figure 2
Plots of the logarithm of the ratio of abundance (cohort $i + 1$ ÷ cohort i) versus sampling period (measured as 5-day intervals) in 1981. Slopes of the linear regression lines estimate $Z_E - Z_L$. (Note that the scale of the ordinate varies among plots.) In 8 of the 11 comparisons of cohorts, the slope is positive (later cohort has the better survival).

ing relative survival. For instance, Crecco and Savoy suggest that cohorts spawned later in the season have a higher relative survival based on comparison of larval and juvenile numbers. Our method provides a tool with which we can test this hypothesis for the larval stage while overcoming problems which may be associated with sampling variability or limited sample sizes.

Ichthyoplankton were sampled throughout the spring of 1980 and 1981. Approximately 30–40 larvae were aged for each 5-day sampling period. We compared the mortality of cohort $i + 1$ with that of cohort i for all possible i where cohort i is defined to be animals born in the ith 5-day period of the season. This resulted in 12 comparisons for 1980 and 11 comparisons for 1981 (Figs. 1, 2). Because information on exact sizes of the samples was not available to us, we computed unweighted regressions. This provides unbiased estimates

of the regression parameters, but the estimates are not of minimum variance and the standard errors are not accurate (Weisberg 1980).

The coefficients of determination were poor (r^2 range 4–98%, mean 55%, for comparisons with three or more observations) suggesting we cannot place much confidence in the magnitude of the estimates of differential mortality ($Z_E - Z_L$). This is undoubtably due to the very small sample sizes. However, it is worth noting that in 12 of the 12 comparisons using the 1980 data, the differential mortality was positive (see Figure 3), i.e., the survival rate of cohort $i + 1$ (the later spawned cohort) was higher than the survival rate of cohort i (the earlier spawned cohort). Also, in 8 out of 11 comparisons using the 1981 data, the later-spawned cohort had the higher survival rate. If there were no differences in survival rates of the two cohorts in each
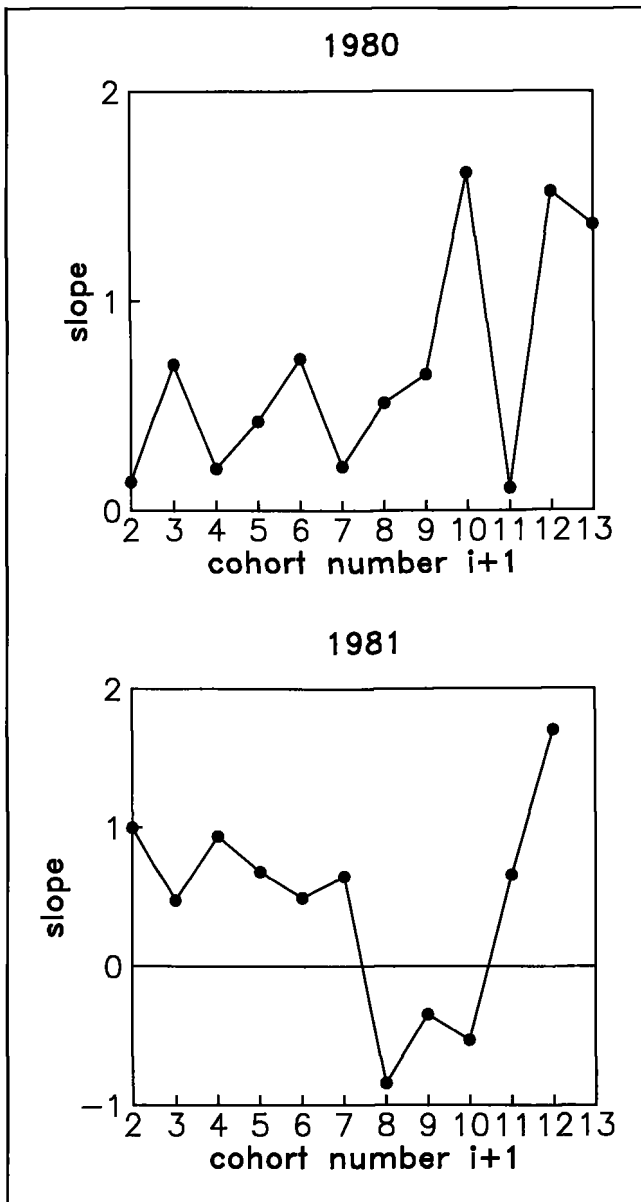
Figure 3

Plot of estimated differential mortality (i.e., Z for cohort $i + 1$ minus Z for cohort i) versus cohort number, $i + 1$. Differential mortalities are estimated as the slopes of the regressions in Figure 1 (top) and Figure 2 (bottom). Note that in all but three cases, the later-spawned cohort is estimated to have lower mortality (higher survival) than the earlier-spawned cohort.

pair, then one would expect half the comparisons would have a positive estimate of differential mortality and half would have negative estimates.

Discussion

We have presented three methods for estimating relative survival rates of larval fishes. The first method is based on a logistic regression model. It provides a graphic way to check the assumptions of constant relative survival and constant ratio of catchabilities. A special case of this method is the two-sample estimator of Paulik and Robson (1969) (our eq. 4). Explicit specification of the likelihood function (eq. 6) is necessary when many small samples are obtained, e.g., from a daily sampling program where the catch per tow is very small.

An intuitive method for estimating relative survival rate would be to estimate the absolute survival rate of each group by the decline in catch-per-unit-effort between two sampling times, and then to take the ratio of the two survival estimates. Thus,

$$\text{alternative estimate of } S_L/S_E = \frac{\dfrac{\hat{C}_{L2}}{\hat{C}_{L1}}}{\dfrac{\hat{C}_{E2}}{\hat{C}_{E1}}} = \frac{\hat{C}_{E1}}{\hat{C}_{E2}}\frac{\hat{C}_{L2}}{\hat{C}_{L1}}. \quad (7)$$

The assumption necessary for the estimation of each survival rate is that the catchability has not changed over time. If one were to obtain an estimate of survival that is unfeasible (>1.0) one would be tempted to discard the data without computing relative survival. However, the expression to the extreme right in (7) is exactly equivalent to the two-sample estimator in (4). Thus, one can validly estimate relative survival rates even when estimates of absolute survival are obtained which are nonsensical. This is because the relative survival estimators do not require the catchabilities to remain constant over time, only the relative catchabilities.

On the basis of existing information, it is not possible to state quantitatively in what proportion of fish populations or in which situations late-spawning larvae will survive better than early-spawning larvae. In our example, sample sizes were quite small ($\sim$30–40 age determinations per 5-day sampling period) so that estimates of relative survival rate were imprecise. It is interesting to note, however, that in 12 of 12 comparisons in the 1980 data and in 8 of 11 comparisons in the 1981 data the later-spawned cohort (week $i + 1$) had a higher survival rate than the earlier-spawned larvae (cohort from week i) (Fig. 3). However, our results indicate that late-hatching (smaller) larvae survive a given calendar period better than early-hatching (larger) larvae which is opposite to the general findings that larval fish mortality rates decrease with increasing size and/or age (Peterson and Wroblewski 1984, McGurk 1986). This suggests that whatever the cause of mortality (e.g., predation, transport), the vulnerability of larval shad in the Connecticut River to this factor increases with age and/or size during the period studied.

Our method was developed for the situation in which cohorts can be sampled simultaneously. Thus, if in a particular week gear efficiency was better than normal, any changes in catchabilities would tend to affect both cohorts so that the ratio of catchabilities might stay relatively constant. It may be tempting to use our method to compare cohorts over the same part of the ontogeny (i.e., cohorts of the same age occurring at different times of the season). However, in this situation there is no reason to believe that variations in catchability over time of the first cohort will be tracked by variations in catchability of the later cohort. Hence, there is no advantage in using our method over traditional methods of estimating absolute survival based on declines in catch-per-unit-effort.

Finally, we note that methods for estimating relative survival have wide applicability beyond the study of larval fishes. A classical problem is Lee's phenomenon in which back-calculated sizes at the first annulus do not agree with the observed sizes of the young fish. Jones (1958) proposed as explanation that faster-growing fish may have a different mortality rate than slower growers. This can be studied by observing the proportion at each age that have a small back-calculated size at the first annulus. Hoenig and Lawing (1983) describe how the logistic model might explain the occurrence of progressively skewed sex ratios with age. Differential survival has been studied by a variety of mark-recapture methods in the context of fitness and natural selection (see Manly 1985) and as a means of estimating impacts of power plants (see Burnham et al. 1987). Another application is to the evaluation of stocking success as related to genetic strain of fish or hatchery treatment. That is, the ratio of abundances of two strains can be monitored over time to study whether one strain survives better than the other.

Acknowledgments

Citations

Burnham, K.P., D.R. Anderson, G.C. White, C. Brownie, and K.H. Pollock
1987 Design and analysis methods for fish survival experiments based on release-recapture. Am. Fish. Soc. Monogr. 5.

Crecco, V., and T. Savoy
1985 Effects of biotic and abiotic factors on growth and relative survival of young American shad, *Alosa sapidissima*, in the Connecticut River. Can. J. Fish. Aquat. Sci. 42:1640–1648.

Crecco, V., T. Savoy, and L. Gunn
1983 Daily mortality rates of larval and juvenile American shad (*Alosa sapidissima*) in the Connecticut River with changes in year-class strength. Can. J. Fish. Aquat. Sci. 40:1719–1728.

Hoenig, J.M., and W.D. Lawing
1983 Using incomplete tagging data to estimate the total mortality rate when the sex ratio is skewed. ICES C.M. 1983/D:24.

Jones, R.
1958 Lee's phenomenon of "apparent change in growth rate", with particular reference to haddock and plaice. Int. Comm. Northwest Atl. Fish., Spec. Publ. 1:229–242.

Manly, B.F.J.
1985 The statistics of natural selection. Chapman and Hall, NY, 484 p.

McCullagh, P., and J.A. Nelder
1983 Generalized linear models. Chapman and Hall, NY, 261 p.

McGurk, M.D.
1986 Natural mortality of marine pelagic fish eggs and larvae: Role of spatial patchiness. Mar. Ecol. Prog. Ser. 34:227–242.

Methot, R.
1983 Seasonal variation in survival of larval northern anchovy, *Engraulis mordax*, estimated from age distribution of juveniles. Fish. Bull., U.S. 81:741–750.

Paulik, G.J., and D.S. Robson
1969 Statistical calculations for change-in-ratio estimators of population parameters. J. Wildl. Manage. 33:1–27.

Peterson, I., and J.S. Wroblewski
1984 Mortality rate of fishes in the pelagic ecosystem. Can. J. Fish. Aquat. Sci. 41:1117–1120.

Rice, J.A., L.B. Crowder, and M.E. Holey
1987 Exploration of mechanisms regulating larval survival in Lake Michigan bloater: a recruitment analysis based on characteristics of individual larvae. Trans. Am. Fish. Soc. 116:703–718.

Rosenberg, A.A., and A.S. Haugen
1982 Individual growth and size-selective mortality of larval turbot reared in enclosures. Mar. Biol. (Berl.) 72:73–77.

Seber, G.A.F.
1982 The estimation of animal abundance and related parameters. Macmillan, NY, 654 p.

Victor, B.C.
1983 Recruitment and population dynamics of coral reef fish. Science (Wash. DC) 219:419–420.

Weisberg, S.
1980 Applied linear regression. John Wiley, NY, 283 p.