

Parameter Estimation for an Egg Production Method of Northern Anchovy Biomass Assessment

SUSAN PICQUELLE and GARY STAUFFER¹

Southwest Fisheries Center,
National Marine Fisheries Service, NOAA,
P.O. Box 271,
La Jolla, CA 92038

ABSTRACT

The estimate of daily egg production is derived from the exponential mortality model fit to the egg density data generated by the plankton survey. The specific formulation of the model used in any particular application depends on the availability of computer programs to run the regression analysis. The model also provides an estimate of egg mortality. Stratification and station weighting schemes depend on survey station pattern. Poststratification of plankton stations is recommended for eliminating those plankton stations occupied that lie beyond the spawning area.

The adult spawning parameters for the northern anchovy example are estimated from trawl survey data using the equations for the sample mean and variance from cluster sampling theory. The choice of the most appropriate estimate in any particular application depends on the sample design. The need for stratifying the survey depends on whether the daily population fecundity parameters change in a consistent fashion over the range and duration of the survey.

The estimation of average weight is straightforward except that the weight of females with hydrated ovaries must be adjusted downward to correct for the temporary weight gain from the increased fluid in the ovaries. Since batch fecundity cannot be measured for each female fish, it is estimated for each individual from a regression model of batch fecundity and ovary-free body weight derived from a sample of female fish with hydrated ovaries. The station value for batch fecundity is estimated as the sample mean of the estimated fecundities. The variance of the mean fecundity, however, is adjusted to include the additional variance resulting from the regression estimation.

For the northern anchovy trawl survey, spawning fraction is the proportion of the mature females which have day-1 postovulatory follicles. Evidence from these trawl surveys indicate female fish classed as day-0 spawners are oversampled. The number of mature females per station is adjusted to compensate for this bias. The sex ratio parameter is the fraction of the mature population that is female, based on weight rather than numbers. Sex ratio data are best generated from a second subsample. For these trawl surveys, the sampling bias of day-0 spawning females occurs during the peak hours of the evening spawning period and impacts the estimates of spawning fraction and sex ratio.

INTRODUCTION

The development of equations for estimating egg production parameters is based on statistical procedures that are applicable to survey and sampling designs and that give valid estimates of the parameters and their variances. The purpose of this paper is to present the statistical equations for estimating daily egg production and daily specific² fecundity parameters and the associated variance estimates that have been used for the northern anchovy. For convenience, a summary of the parameter values estimated for the northern anchovy from 1980 to 1984 is included in Table 1.

Table 1.—Time series of egg production parameters (1980-84).

Parameters		1980	1981	1982	1983	1984
Daily egg production (10 ¹² eggs/d)	<i>PA</i>	26.34	20.96	13.51	17.25	12.98
Average female weight (g)	<i>W</i>	17.44	13.37	18.83	11.20	12.02
Batch fecundity (no. eggs/batch per mature female)	<i>F</i>	7,751	8,329	10,845	5,297	5,485
Spawning fraction (no. spawning females per mature female)	<i>S</i>	0.142	0.106	0.120	0.094	0.160
Sex ratio (no. females/total)	<i>R</i>	0.478	0.501	0.472	0.549	0.582
Daily specific fecundity (no. eggs/g biomass per d)		30.28	33.03	32.53	24.35	42.43
Spawning biomass (10 ³ t)	<i>B</i>	870	635	415	652	309

DAILY EGG PRODUCTION

The estimate of daily production of eggs released into the sea, P , by spawning adult fish is derived by regressing the counts of eggs on their age using the exponential mortality model. This model assumes a constant mortality rate.

$$P_{ijk} = P_i e^{-Z t_{ijk}} + \epsilon_{ijk} \quad (1)$$

where P_{ijk} = the number of eggs in day k age category from station j in stratum i ,

t_{ijk} = the age in days measured as the elapsed time from the specified spawning time, t_0 , to the time of sampling of station j in stratum i ,

P_i = the daily egg production per unit area in stratum i ,

Z = the daily rate of instantaneous egg mortality, and

ϵ_{ijk} = the additive error term.

A number of regression procedures can be used to estimate P and Z from the observations of P_{ijk} and t_{ijk} depending on the design of the survey and the availability of statistical computer programs.

First, the need to stratify the plankton survey should be evaluated. Stratification is usually undertaken to reduce the variance of parameter estimates. If more than one major spawning area exists, or if the survey is conducted over a relatively long period such that spawning rates differ among areas or over time, stratification of the

¹Present addresses: Northwest and Alaska Fisheries Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115.

²Daily specific fecundity = number of eggs produced/gram of biomass per day.

data after the survey would be appropriate. This occurred in the anchovy spawning survey conducted in 1983 (Picquelle and Hewitt 1984).

In addition, the total survey area may be considerably larger than the spawning range. This has been the case for the anchovy resource off the California coast. The anchovy surveys are intentionally designed to sample the full range of anchovy spawning, knowing that in any one year the spawning will occur over a much smaller area within the range. As a result, many stations are beyond the spawning area of anchovies and contribute a large number of stations with zero egg counts, thus inflating the variance of the egg count data set. To reduce the impact of these zero-count stations, the total survey area is poststratified into two strata depending on the presence or absence of eggs in the sample. Stratum 0 contains the stations beyond the geographic area of spawning so that P is zero for stratum 0. The geographic area containing all the anchovy eggs and the few embedded stations with zero egg counts make up stratum 1. In this latter case, the size of the geographic area of stratum 1, A_1 , and the number of occupied stations, n_1 , become random variables with variances and expected values. The magnitude of this added variance is negligible if n_1 is larger than 100 (Jessen 1978). Poststratification of the stations to eliminate those stations beyond the spawning area is a recommended procedure for estimating egg production.

In addition to stratification, it is desirable to increase the density of plankton stations within the geographic area where adult spawning is expected. Increased sample size in this area will reduce the variance of P for the positive stratum. To correct for differences in station density, the egg count observations by station need to be weighted by a factor proportional to their representative area such that the sum of the station weights in the positive stratum, $i=1$, equals n_1 , the total number of stations in this stratum, i.e.,

$$\sum_{j=1}^{n_1} w_{1j} = n_1.$$

Since the egg counts in the zero stratum are zero, weighting in that stratum is unnecessary. The preferred statistical technique for estimating parameters P and Z from Equation (1) is weighted nonlinear least squares regression fit to individual egg counts and ages from stations within the positive stratum. The computer program used for the northern anchovy case was BMDPAR (Dixon and Brown 1981).

The stratified estimate of P can be calculated as the weighted average of P_0 and P_1 , where P_0 is zero by definition and P_1 is estimated by regression analysis, and the weights are the relative areas of the two stratum, i.e.,

$$P = \frac{A_1}{A} P_1 + \frac{A_0}{A} P_0$$

and the variance, adjusted for postsurvey stratification (Jessen 1978) is

$$Var(P) = (1 + \frac{1}{n}) (\frac{A_1}{A} Var(P_1) + \frac{A_0}{A} Var(P_0))$$

where A_i = the area of stratum i for each region,
 A = $A_1 + A_0$,
 n = the total number of observations for the survey,
 $Var(P_1)$ = estimated for stratum 1 from the regression analysis, and
 $Var(P_0) = 0$ by definition.

In the event that a computer program for nonlinear regression is not available, the egg mortality model must be linearized so that P and Z can be estimated by linear regression methods. The linear

version of the model is

$$\ln(P_{ijk}) = \ln P_1 - Zt_{ijk} + \epsilon_{ijk}.$$

This linear model gives the error structure of ϵ as additive in the transformed expression. This implies a multiplicative error structure in the nonlinear model. The form of the error in either case should be studied by examining the variability about P_{ijk} versus age. This can best be accomplished by examining the residuals between the observed P_{ijk} and those predicted by the model. The antilog estimate of P will be biased. An unbiased estimate can be approximated by

$$P = e^{(\ln P + s^2/2)}$$

where s^2 is the estimated residual variance from the regression analysis, usually denoted by $S_{y \cdot x}^2$.

If computer facilities are not readily available, the estimation of P and Z can be further simplified. This simplification involves averaging the egg counts P_{ijk} over time intervals of equal length. A minimum of three time intervals is necessary to estimate the two parameters P and Z . For the anchovy example, five intervals of 12 h were tried. The modified mortality model in this case is³

$$\bar{P}_{ik} = P_1 e^{-Zt'_k} (1 - e^{-Z\Delta t}) / Z\Delta t$$

where \bar{P}_{ik} = the average number of eggs of age t_k sampled during the time interval (t_k, t_{k+1}) ,
 $t'_k = t_k - t_0$ = the time elapsed (or age) between spawning and the beginning of the time interval, k ,
 t_0 = the midpoint of the daily spawning period, and
 Δt = the length of the time interval $(t_{k+1} - t_k)$ over which P_{ijk} are averaged.

The linear form of this model is

$$\ln(P_{ik}) = \ln(P_1(1 - e^{-Z\Delta t}) / Z\Delta t) - Zt'_k$$

where the regression coefficients $b_0 = \ln(P_1(1 - e^{-Z\Delta t}) / Z\Delta t)$, and $b_1 = -Z$. Substituting the value of Z from b_1 into b_0 will provide an estimate of P_1 . These estimates of P and Z should be useable in most cases, but their associated variance estimates will not be realistic.

DAILY SPECIFIC FECUNDITY

Station Weighting Alternatives

In sampling theory, several estimates of the population mean are described; the choice of the most appropriate estimate depends on the sample design. Trawl surveys typically consist of a three-stage sampling design: 1) the selection of the trawl stations, 2) the catch of fish at the station, and 3) the selection of the subsample of fish from the catch. The sample design, and hence the estimate, is determined by how each of the stages is executed.

There are two common methods for selecting station locations: random sampling and judgment sampling. Under the random sampling regime, the trawl stations are distributed randomly over the survey area. The exact locations of the stations are determined principally by the requirements of convenient and efficient use of ship time, and by the desire to have stations distributed evenly over the

³McCaughan, C. A., Intl. Pacific Halibut Comm., Seattle, WA., pers. commun. May 1981.

survey area. This strategy will produce station locations independent of local fish population densities.

Judgment sampling uses information independent of the trawl samples to place stations where fish abundances are high. This information may consist of a historical account of the distribution of fish based on current sea temperature and salinity data, results from a recent survey, the use of sonar to detect fish schools, or observable evidence of local fish concentrations, such as their spawning products in plankton samples. The resulting distribution of stations will be patchy, with the high densities of stations coinciding with high densities of fish. This strategy approximates the sampling technique of probability proportional to a measure of size.

The second stage of sampling, the catch of fish, is ideally a random sample of fish residing at the station. However, some bias may occur due to net design, the execution of the trawl, fish behavior, and other variables. In addition to striving for a random sample, it is also advantageous if the catch size is proportional to the number of fish at the station. This is valuable information which may be incorporated into the estimate, but this situation holds only for certain species and sampling methods.

Fish are subsequently sampled randomly from the catch. The subsample size is usually constant, but may vary with the catch size to produce a self-weighting estimate for some of the estimates presented below.

Based on this sampling structure, many specific sample designs and corresponding estimates of the population mean may be considered. The following estimates all assume that the total population size is very large and the finite population correction is approximately 1. Under these conditions, the within-subsample contribution to variance disappears because this quantity is multiplied by $1/Nn$ which is approximately 0.

If stations are selected randomly, and catch size is unrelated to fish abundance, then each station should receive equal weight and equal subsample sizes should be attempted. In this case the appropriate estimate is (Cochran 1977)

$$\bar{y} = \frac{\sum_{i=1}^n \bar{y}_i}{n} \quad \text{and} \quad \hat{V}ar(\bar{y}) = \frac{\sum_{i=1}^n (\bar{y}_i - \bar{y})^2}{n(n-1)} \quad (2)$$

where \bar{y} = the estimate of the population mean,
 n = the number of stations,

$\bar{y}_i = \frac{\sum_{j=1}^{m_i} y_{ij}}{m_i}$ = the mean of the i th station, and

m_i = the number of fish subsampled from the i th catch.

This estimate is a biased estimate of the true population mean and this bias does not necessarily get small as n gets large. The estimate is self-weighting if m_i is constant.

A better estimate exists for the case of random station selection if the catch size is proportional to the abundance of fish at the station. This is the ratio-to-size estimate:

$$\bar{y} = \frac{\sum_{i=1}^n M_i \bar{y}_i}{\sum_{i=1}^n M_i} \quad \text{and} \quad \hat{V}ar(\bar{y}) = \frac{\sum_{i=1}^n M_i^2 (\bar{y}_i - \bar{y})^2}{\left[\sum_{i=1}^n \frac{M_i}{n} \right]^2 n(n-1)} \quad (3)$$

where M_i is the total number of the target fish species caught in the i th trawl. This estimate is biased because it is a ratio of two random variables, but the bias is small and gets smaller as n gets

larger. This estimate is self-weighting when the subsampling fraction m_i/M_i is constant.

Under judgment sampling, the attempt is made to sample with probability proportional to size. If this can be accomplished, or if confidence is high that this situation is closely approximated, then the unbiased estimates of the population mean and the variance are the same as Equation (2). This estimate is self-weighting when m_i is constant.

If, instead, the stations are selected with probability proportional to u_i , a measure of size, and the catch size is proportional to the population size at the station, then the appropriate estimate is

$$\bar{y} = \frac{1}{M_0 n} \sum_{i=1}^n \frac{M_i \bar{y}_i}{u_i} \quad \text{and} \quad \hat{V}ar(\bar{y}) = \frac{\sum_{i=1}^n \left[\frac{M_i \bar{y}_i}{u_i M_0} - \bar{y} \right]^2}{n(n-1)} \quad (4)$$

where M_0 is the total population size. This estimate is self-weighting if $nu_i m_i / M_i$ is constant. However, this estimate is not very useful because M_0 , the total population size, is rarely known, and the estimates of station sizes, u_i , can rarely be enumerated.

A modification of Equation (2) is used for the trawl survey for the northern anchovy. As mentioned earlier, this estimate is unbiased if sampling with probability proportional to size. The information used to detect high concentrations of anchovies is the occurrence of anchovy spawning products in the plankton samples taken concurrently, and the presence of apparent schools on the sonar. Both of these factors are good indicators of local concentrations of anchovies, and the resulting sample design is assumed to be a good approximation of sampling with probability proportional to size, justifying the use of Equation (2).

Equal subsample sizes are attempted, but occasionally a station will produce a very small catch or a catch with very few mature females (mostly males or mostly immature fish). Both situations will result in a small subsample, as most of the parameters to be estimated are for mature females. This occurrence is interpreted as meaning that an error in judgment sampling has been made. The actual size of the station, based on mature females, is much smaller than was estimated at the time the trawl station was selected. Hence, the probability of selecting that station should be adjusted *a posteriori* to reflect the actual size of the station. This is accomplished by giving these stations less weight in the estimate; each station is weighted by its subsample size, m_i . Thus, Equation (2) is modified to produce the following estimate:

$$\bar{y} = \frac{\sum_{i=1}^n m_i \bar{y}_i}{\sum_{i=1}^n m_i} \quad \text{and} \quad \hat{V}ar(\bar{y}) = \frac{\sum_{i=1}^n m_i^2 (\bar{y}_i - \bar{y})^2}{\left[\sum_{i=1}^n \frac{m_i}{n} \right]^2 n(n-1)} \quad (5)$$

If m_i is constant, this estimate simplifies to Equation (2). This estimate is biased because it is the ratio of two random variables, $m_i \bar{y}_i$ and m_i . However, the bias is of the order $1/n$, so that the bias gets smaller as n gets larger.

Trawl Survey Stratification

The egg production model assumes that the parameters in the model are constant over the range and duration of the survey. If this assumption is violated, the survey should be divided into regions or time spans within which the parameters are constant. The biomass is then estimated separately for each section of the survey and then summed to produce the total biomass estimate.

An example of this situation is the Spring 1983 anchovy spawning biomass survey. Two parameters of the daily specific fecundity varied significantly with geographic regions, female weight decreased from north to south, and spawning fraction increased from north to south.

The survey area was divided into three regions: north, bight, and south (Figs. 1, 2). Figure 3 illustrates the frequency distribution for female weight by region; the average female weight for the southern region was significantly smaller than the average weights for both the bight and north regions. The pattern for spawning fraction was the opposite (Fig. 4) with the estimate for the north region being significantly smaller than the estimates for the bight and south regions.

Table 2 presents the parameter estimates and biomass estimates for each region. Each of the population fecundity estimates, W , S , F , and R , was estimated separately for each region. It was impossible to estimate P_0 , egg production, independently for each region because of the small sample sizes per region. Instead, the slope of the mortality curve, Z , was assumed to be constant for the entire survey and hence was estimated using all the data. The intercept of the mortality curve, P_0 , was allowed to vary between regions and was estimated by fitting a separate but parallel mortality curve to each region while holding the mortality estimate, Z , fixed at the value previously estimated.

The total biomass estimate is simply the sum of the regional biomass estimates, and the total variance estimate is also just the sum of the regional variances. This variance estimate ignores any covariance terms between regions, which is probably trivially small because all parameters were estimated using separate and independent data for each region, except for the mortality, Z , whose contribution to covariance is probably slight.

Parameters		North	Bight	South	Total
Daily egg production (no. eggs/0.05 m ² per d)	P	1.62 (0.671)	7.28 (0.0751)	5.06 (0.332)	
Area of region (10 ¹² 0.05 m ²)	A	0.420	1.33	1.36	
Average female weight (g)	W	12.9 (0.121)	11.2 (0.0705)	9.63 (0.0385)	
Batch fecundity (no. eggs/batch per mature female)	F	6,285 (0.140)	5,295 (0.0882)	4,423 (0.0570)	
Spawning fraction (no. spawning females per mature female)	S	0.0346 (0.563)	0.103 (0.174)	0.126 (0.237)	
Sex ratio (no. females/total)	R	0.523 (0.0949)	0.559 (0.0736)	0.549 (0.128)	
Spawning biomass (10 ³ t)	B	77.5 (0.897)	358 (0.214)	216 (0.419)	652 (0.211)

Parameter Estimation

The parameters of the daily population fecundity are all estimated from samples of anchovies collected on a midwater trawl survey. These parameters, female weight (W), batch fecundity (F), spawning fraction (S), and sex ratio (R), and their variances, are estimated using Equation (5) developed previously:

$$\bar{y} = \frac{\sum_{i=1}^n m_i \bar{y}_i}{\sum_{i=1}^n m_i} \quad \text{and} \quad \widehat{Var}(\bar{y}) = \frac{\sum_{i=1}^n m_i^2 (\bar{y}_i - \bar{y})^2}{\left(\sum_{i=1}^n \frac{m_i}{n}\right)^2 n(n-1)}$$

where n = the number of trawls,
 m_i = the number of fish subsampled from each trawl,

$$\bar{y}_i = \frac{\sum_{j=1}^{m_i} y_{ij}}{m_i} = \text{the average value for the } i\text{th trawl, and}$$

y_{ij} = the observed value for the j th fish sampled from the i th trawl.

Female weight is estimated from a fixed subsample size of mature females. The subsample size has ranged from 15 to 25 mature females for the surveys taken from 1980 through 1983; however, the targeted subsample size is not always realized, due to very small catches, a high proportion of immature fish, or a high proportion of males. The y_{ij} in Equation (5) is the whole body weight of the j th mature females sampled from the i th trawl (W_{ij}). This observed weight is adjusted downward for those females whose ovaries contain hydrated eggs because their body weight is temporarily inflated due to water retention. This adjusted weight (\hat{W}_{ij}) is estimated from a linear regression of whole body weight regressed on ovary-free weight (W_{ij}^*) which is fit only to those females that do not have hydrated eggs.

$$\hat{W}_{ij} = \hat{\alpha} + \hat{\beta} W_{ij}^* \quad (6)$$

The observed frequency distribution of the average female weight per trawl is usually symmetrical although there may occur a hint of bimodality if there is a large 1-yr-old year class. The weights within each trawl tend to be homogeneous, suggesting that the anchovy schools are homogeneous with regard to weight.

Batch fecundity can be observed only for those females whose ovaries contain hydrated eggs. There is a high correlation between the number of eggs per batch and the ovary-free body weight. This relationship is used to estimate batch fecundity for the same mature females used to estimate female weight.

The sample of hydrated females is collected throughout the trawl survey and the number of eggs per batch (F_{ij}) and ovary-free weight (W_{ij}^*) are recorded for each of these females. This data are used to fit a model regressing batch fecundity on ovary-free weight

$$F_{ij} = \alpha + \beta W_{ij}^* + \epsilon_{ij} \quad (7)$$

A linear regression has explained the data satisfactorily in previous surveys, although a curvilinear model should be considered depending on the shape of the data.

Using this regression, batch fecundity is estimated for each mature female subsampled. The accuracy of the estimated batch fecundity will be improved if the distribution of weights for the sample of hydrated females used to fit the regression is similar to the distribution of weights for the total sample of females.

Equation (5) is used to estimate average batch fecundity where the y_{ij} are in this case the estimates \hat{F}_{ij} . However, there is an added source of variance that should be included because the \hat{F}_{ij} are not observed directly but are estimates with their own associated variance. Thus, the estimate of variance is adjusted to include this additional variance (Draper and Smith 1966):

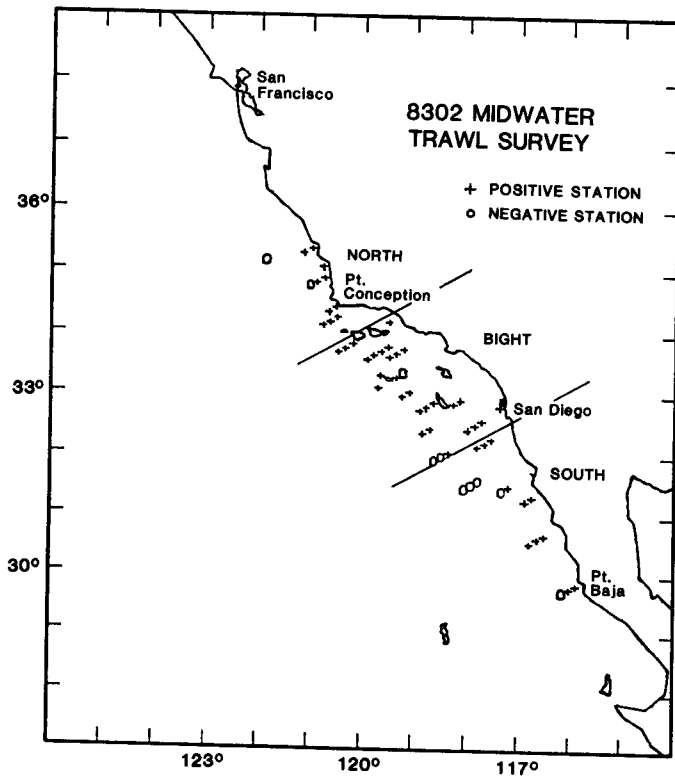


Figure 1.—Geographic distribution of trawl stations and positive trawls within each region.

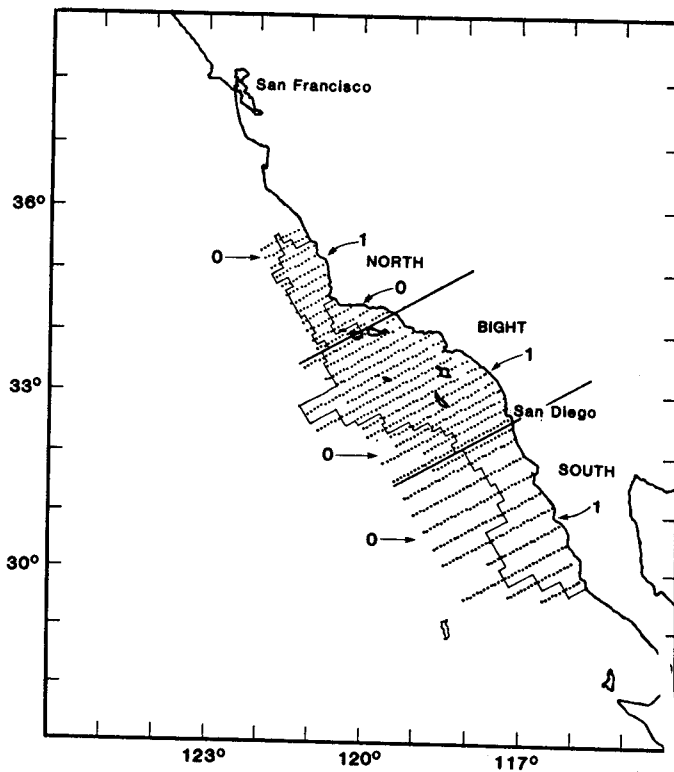


Figure 2.—Subdivision of CalVET survey into strata (0 = beyond the range of anchovy spawning; 1 = within the range of anchovy spawning) and regions (North, Bight, and South).

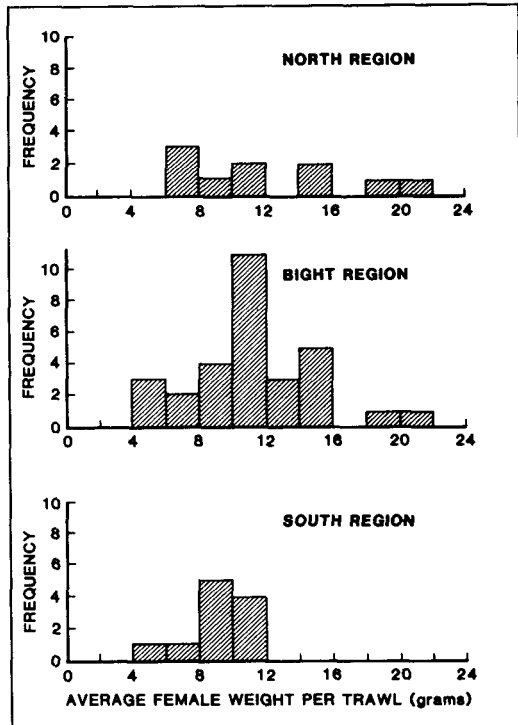


Figure 3.—Frequency distribution for average whole-body weight of mature females in grams for each region.

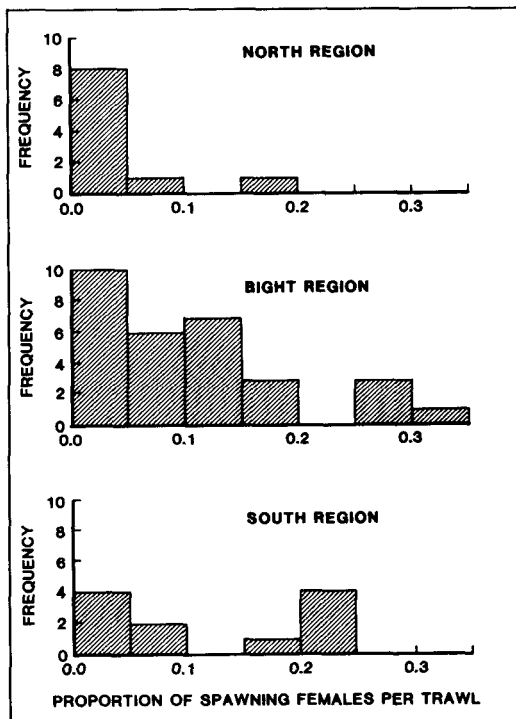


Figure 4.—Frequency distributions of spawning fractions for each region.

$$\hat{V}ar(\bar{F}) = \frac{\sum_{i=1}^n m_i^2 \left[\frac{(\bar{F}_i - \bar{F})^2}{n-1} + \frac{s_h^2}{n_h} + (\bar{W}_i^* - \bar{W}_h^*) \hat{V}ar(\hat{\beta}) \right]}{\left(\sum_{i=1}^n \frac{m_i}{n} \right)^2 n} \quad (8)$$

where \bar{F} = the estimate of batch fecundity for the whole population of mature females,
 \bar{F}_i = the average batch fecundity for the i th trawl,
 \bar{F}_j = $\sum_{j=1}^{m_i} \hat{F}_{ij} / m_i$ where \hat{F}_{ij} is the estimated batch fecundity for the j th female in the i th subsample,
 s_h^2 = the variance about the regression (Equation (7)),
 n_h = the number of hydrated females used to fit the regression (Equation (7)),
 \bar{W}_i^* = the average ovary-free weight for the i th trawl,
 \bar{W}_h^* = the average ovary-free weight for the n_h hydrated females, and
 $\hat{V}ar(\hat{\beta})$ = the variance of the slope of the regression (Equation (7)).

Spawning fraction is estimated using Equation (5) by setting \bar{y}_i equal to the proportion of mature females in the i th trawl which have been classified as having day-1 postovulatory follicles (day-1 spawners). Thus, \bar{y}_i estimates the fraction of mature females in the population that are day-1 spawners, which is a measure of the fraction of mature females which are spawning on any given day.

Spawning fraction can also be estimated by the fraction of females classified as day-0 spawners. However, this fraction has been consistently higher than the fraction of day-1 spawners. Past experience has shown that using day-0 spawners may produce a biased estimate, at least for northern anchovy sampled by a trawl survey. Evidence of this conclusion will be presented in a later section.

If day-0 spawners are indeed oversampled, then day-1 spawners are undersampled. The sample of mature females from each catch may be grouped into three categories: Day-0 spawners (m_i^0); day-1 spawners (m_i^1); and day 2+ spawners (m_i^2). For a fixed subsample size, if one group is over-represented then the other two groups are under-represented. This is corrected by adjusting m_i to reflect what subsample size would have included the observed number of day-1 spawners if day-0 had not been oversampled. The number of day-0 spawners included in m_i is replaced by the observed number of day-1 spawners, since day-0 and day-1 spawners should be equal, on average, because they both measure the number of females spawning during a 24-h period. Thus, the m_i in Equation (5) is replaced by

$$m_i^* = 2m_i^1 + m_i^2. \quad (9)$$

The average value of m_i^* will be smaller than the average m_i , and the resulting estimate of spawning fraction based on day-1 spawners will be larger to compensate for the bias in sampling day-0 spawners.

The parameter sex ratio is the fraction of the mature population that is female, based on weight rather than numbers. Equation (5) is again used, where m_i is the weight of the subsample rather than number, and \bar{y}_i is the fraction of the subsample weight that is attributable to female fish. Both mature and immature fish are included in the estimate because it is impossible to distinguish between mature and immature males. It is assumed that sex ratio by weight is the same for both mature and immature fish.

To save effort in preserving and weighing individual fish, \bar{y}_i and m_i are estimated rather than measured directly. A fixed number of fish are subsampled from each trawl and the numbers of females and males are recorded. The average weight for each sex is estimated

for each trawl from a smaller fixed subsample of each sex. The total weight of each sex in the subsample is estimated by multiplying the observed number of fish of that sex by its average weight.

$$\begin{aligned}\hat{W}_i^F &= m_i^F \cdot \bar{W}_i^F \\ \hat{W}_i^M &= m_i^M \cdot \bar{W}_i^M\end{aligned}\quad (10)$$

where \hat{W}_i^k = the total estimated weight of the k th sex in the i th subsample,

m_i^k = the number of fish of the k th sex in the i th subsample, and

\bar{W}_i^k = the average weight of the k th sex in the i th subsample.

Then m_i is estimated by the sum of the estimated total weight of males plus the estimated total weight of females,

$$\hat{m}_i = \hat{W}_i^F + \hat{W}_i^M \quad (11)$$

and \bar{y}_i is the estimated total weight of females divided by m_i ,

$$\bar{y}_i = \hat{W}_i^F / \hat{m}_i \quad (12)$$

The parameters W_i^F and W_i^M are estimated with little error because the weights of fish within a trawl catch are quite homogeneous. Thus the added variance in estimating sex ratio, due to the fact that \bar{y}_i and m_i are estimated rather than observed, is assumed to be trivial.

Sampling Bias

Spawning fraction may be estimated by either the number of day-0 spawners or the number of day-1 spawners, as each is an estimate of the number of females spawning on any given day. However, as mentioned earlier, the proportion of day-0 spawners is consistently larger than the proportion of day-1 spawners. A hypothesis on the mechanism causing this result is that females who are actively spawning are more vulnerable to capture by a midwater trawl. This hypothesis is supported by the observation that the catch of spawning females (day-0) increases significantly during the hours of the evening when spawning takes place.

This phenomenon is illustrated by Figure 5. This bar chart was calculated using data from the trawl surveys conducted from 1978 to 1980. Day-0 spawners are composed of three types of females: Females whose ovaries contain hydrated eggs which will be spawned later that evening (represented by an open bar); females whose ovaries contain postovulatory follicles indicating they have spawned earlier that evening (represented by the bar area with horizontal lines); and females whose ovaries contain both hydrated eggs and postovulatory follicles indicating they were caught during spawning (represented by vertical lines in the bars). The predominant hours of spawning are 2100-2359, when the highest number of females are caught with both hydrated eggs and postovulatory follicles in their ovaries. This is also the time period when the proportion of day-0 spawners in the catch is highest.

It should also be noted that the proportion of day-0 spawners decreases after 0300. This can be more easily seen in Figure 6, which shows the proportion of day-0 spawners by hour for each year, and all years combined. Figure 7 shows that the proportion of day-1 spawners also drops off after 0300. This is explained by an increased error rate in classifying ovaries as to the date of spawning (day-0, day-1, or later), and in subsequent years trawl catches were not made after 0300 to avoid this problem.

If one ignores the points corresponding to 0300 and later, then

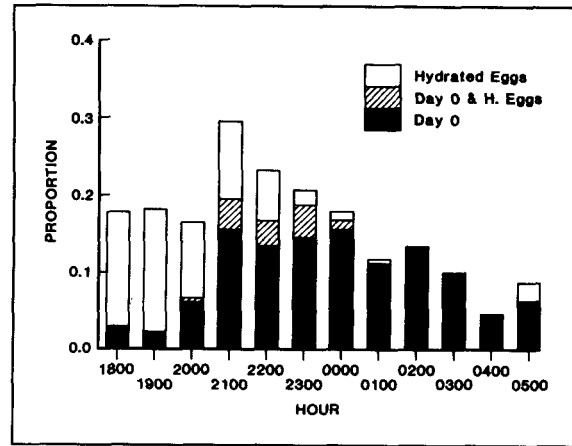


Figure 5.—Breakdown of day-0 females into ovarian categories by time of sampling.

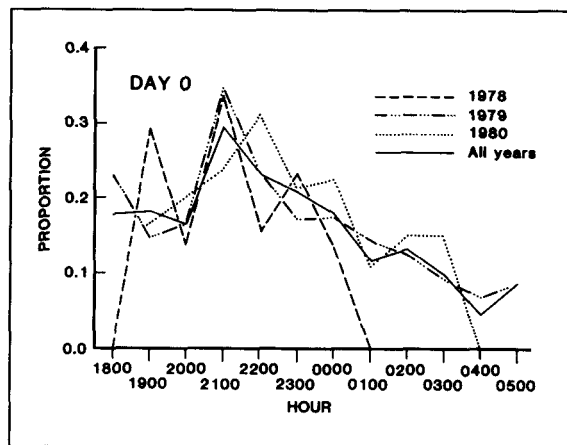


Figure 6.—Proportion of day-0 spawning females by hour for trawl surveys conducted from 1978 to 1980.

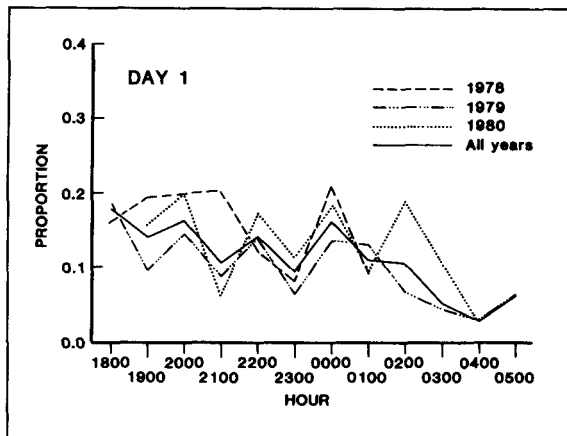


Figure 7.—Proportion of day-1 spawning females by hour for trawl surveys conducted from 1978 to 1980.

a comparison of Figures 6 and 7 shows that the plot of day-1 spawners is quite flat over time, while day-0 spawners exhibit a definite peak at 2100-2359. The comparison is highlighted in Figure 8. The proportion of day-0 females is plotted against the proportion of day-1 females for each hour, and 1-standard-error bars are drawn for day-1 proportions. The diagonal line shows the values for which the proportions of day-0 and day-1 spawners are equal. The obvious outlier points (i.e., those points furthest from the day 0=day 1 line) correspond to the time 2100-2359.

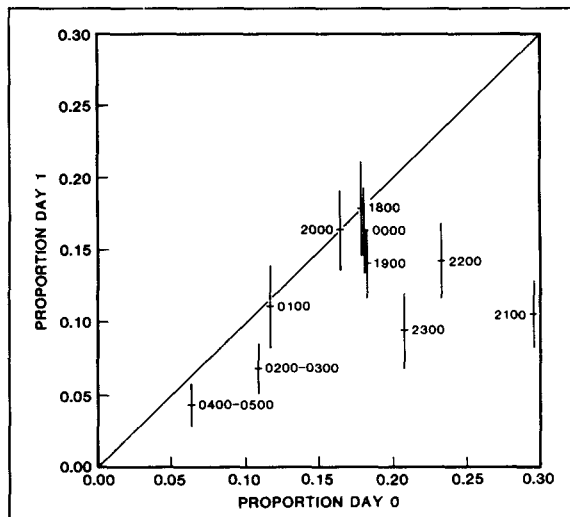


Figure 8.—Day-1 vs. day-0 females for each hook with standard error bars for day-1 females.

The effect of time on the proportion of day-0 and day-1 females in the catch can be quantified using contingency table analysis. Table 3 shows the number of females from the 1978-80 trawl survey samples, by spawning condition and time of their capture. Parentheses contain the expected number of females under the hypothesis that there is no interaction between time and spawning condition. The contribution to the χ^2 statistic by the cell is in brackets. The resulting χ^2 statistic ($\chi^2 = 38.85$, $df = 16$) is highly significant ($P = 0.05$), thus rejecting the hypothesis of no interaction between time and spawning condition. Examination of the individual cell's contribution to the test statistic shows that the largest deviations are due to the hour 2100-2159. Omitting this hour from the analysis produces Table 4 and a nonsignificant ($P = 0.05$) test statistic ($\chi^2 = 20.51$, $df = 14$). Therefore, if the hour 2100-2159 is omitted, there is no significant relationship between time and spawning condition. Based on this analysis, the conclusion is made that day-0 spawners are sampled with bias during the peak hours of spawning.

A similar scenario has been discovered for sex ratio and time. The proportion of females declines radically during the time period 2300-2359. A contingency table analysis (Table 5) shows that there is a significant interaction between time and sex ratio ($\chi^2 = 126.44$, $df = 8$); an examination of the individual cells shows a large deviation during time 2300-2359. When this time period is removed (Table 6), there is no significant interaction ($\chi^2 = 10.63$, $df = 7$). Apparently, there is a sample bias also occurring for sex ratio during the hours when spawning activity is greatest.

LITERATURE CITED

COCHRAN, W. G.
1977. Sampling techniques. 3rd ed. John Wiley and Sons, N.Y., 428 p.
DIXON, W. J., and M. B. BROWN.
1981. BMDP statistical software. Univ. Calif. Press, Berkeley, CA., 735 p.
DRAPER, N. R., and H. SMITH.
1966. Applied regression analysis. John Wiley & Sons, N.Y., 407 p.
JESSEN, R. J.
1978. Statistical survey techniques. John Wiley & Sons, N.Y., 520 p.
PICQUELLE, S. J., and R. P. HEWITT.
1984. Estimates of the spawning biomass of the northern anchovy for the 1983-84 California fishing season. Calif. Coop. Oceanic Fish. Invest. Rep. 25:16-27.

Table 3.—Contingency table analysis of spawning condition by time. Each cell contains the observed count, the expected count in parentheses, and the contribution to the χ^2 statistic in brackets.

Spawning condition	Time (h)									Total
	1800-1859	1900-1959	2000-2059	2100-2159	2200-2259	2300-2359	0000-0059	0100-0159	0200-0259	
Day-0	25 (27.5) [0.22]	42 (45.3) [0.24]	30 (35.9) [0.96]	67 (44.5) [11.36]	49 (41.4) [1.41]	30 (28.4) [0.09]	29 (31.8) [0.24]	19 (32.0) [5.26]	9 (13.3) [1.41]	300
Day-1	25 (17.8) [2.87]	31 (29.4) [0.08]	30 (23.3) [1.91]	19 (28.9) [3.41]	27 (26.9) [0.00]	12 (18.5) [2.27]	26 (20.6) [1.39]	18 (20.8) [0.37]	7 (8.7) [0.32]	195
Day-2+	90 (94.7) [0.23]	158 (156.3) [0.02]	23 (123.8) [0.01]	141 (153.6) [1.03]	135 (142.7) [0.42]	103 (98.1) [0.25]	107 (109.6) [0.06]	126 (110.3) [2.25]	52 (46.0) [0.78]	1,035
Total	140	231	183	227	211	145	162	163	68	1,530

$\chi^2 = 38.85$, $df = 16$
 $\chi^2_{16}(0.95) = 26.3$

Table 4.—Contingency table analysis of spawning condition by time, in which the hour 2100-2159 has been omitted. Each cell contains the observed count, the expected count in parentheses, and the contribution to the χ^2 statistic in brackets.

Spawning condition	Time (h)								Total
	1800-1859	1900-1959	2000-2059	2200-2259	2300-2359	0000-0059	0100-0159	0200-0259	
Day-0	25 (25.0) [0.00]	42 (41.3) [0.01]	30 (32.7) [0.23]	49 (37.7) [3.37]	30 (25.9) [0.64]	29 (29.0) [0.00]	19 (29.1) [3.53]	9 (12.2) [0.82]	233
Day-1	25 (18.9) [1.96]	31 (31.2) [0.00]	30 (24.7) [1.13]	27 (28.5) [0.08]	12 (19.6) [2.94]	26 (21.9) [0.78]	18 (22.0) [0.73]	7 (9.2) [0.52]	176
Day-2+	90 (96.1) [0.38]	158 (158.5) [0.00]	123 (125.6) [0.05]	135 (144.8) [0.66]	103 (99.5) [0.12]	107 (111.1) [0.15]	107 (111.8) [1.79]	52 (46.7) [0.61]	894
Total	140	231	183	211	145	162	163	68	1,303

$\chi^2 = 20.51, df = 14$
 $\chi^2_{14}(0.95) = 23.7$

Table 5.—Contingency table analysis of sex ratio by time. Each cell contains the observed count, the expected count in parentheses, and the contribution to the χ^2 statistic in brackets.

Sex	Time (h)									Total
	1800-1859	1900-1959	2000-2059	2100-2159	2200-2259	2300-2359	0000-0059	0100-0159	0200-0259	
Female	153 (131.2) [3.62]	349 (331.2) [0.96]	189 (160.3) [5.13]	271 (268.3) [0.03]	227 (198.7) [4.03]	189 (321.9) [54.88]	163 (154.4) [0.48]	187 (175.5) [0.75]	71 (57.4) [3.23]	1,799
Male	158 (179.8) [2.64]	436 (453.8) [0.70]	191 (219.7) [3.74]	365 (367.7) [0.02]	244 (272.3) [2.94]	574 (441.1) [40.05]	203 (211.6) [0.35]	229 (240.5) [0.55]	65 (78.6) [2.36]	2,465
Total	311	785	380	636	471	763	366	416	136	4,264

$\chi^2 = 126.44, df = 8$
 $\chi^2_8(0.95) = 15.5$

Table 6.—Contingency table analysis of sex ratio by time, in which the hour 2300-2359 has been omitted. Each cell contains the observed count, the expected count in parentheses, and the contribution to the χ^2 statistic in brackets.

Sex	Time (h)								Total
	1800-1859	1900-1959	2000-2059	2100-2159	2200-2259	0000-0059	0100-0159	0200-0259	
Female	153 (143.0) [0.70]	349 (361.0) [0.40]	189 (174.8) [1.16]	271 (292.5) [1.58]	227 (216.6) [0.50]	163 (168.3) [0.17]	187 (191.3) [0.10]	71 (62.5) [1.14]	1,610
Male	158 (168.0) [0.59]	436 (424.0) [0.34]	191 (205.2) [0.99]	365 (343.5) [1.34]	244 (254.4) [0.43]	203 (197.7) [0.14]	229 (224.7) [0.08]	65 (73.5) [0.97]	1,891
Total	311	785	380	636	471	366	416	136	3,501

$\chi^2 = 10.63, df = 7$
 $\chi^2_7(0.95) = 14.1$



UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE

Southwest Fisheries Center
P.O. Box 271
La Jolla, California 92038

April 16, 1987

MEMORANDUM FOR: USERS OF THE EGG PRODUCTION METHOD FOR
ESTIMATING SPAWNING BIOMASS OF PELAGIC FISH.

FROM: REUBEN LASKER *Reuben Lasker*

SUBJECT: ERRATA FOR NOAA TECHNICAL REPORT NMFS 36; "AN EGG
PRODUCTION METHOD FOR ESTIMATING SPAWNING BIOMASS OF PELAGIC
FISH: APPLICATION TO THE NORTHERN ANCHOVY".

A number of printing errors have been discovered by Dr.
Sachiko Tsuji in the published account of the egg production
method. These are important and warrant this memo. Please make
these corrections in your copy.

p. 5, Abstract, 4th line should read:

"be estimable and spawning rate constant over the field
sampling interval."

p. 12, in equation 8, $\hat{\beta}$ should be β .

p. 17, Table 1. on the January line +3.5 should be -3.5.

p. 20, two lines under the formula in the second column,
"sample size" should be "sample scale" and σ_i should read σ_i^2 .
Five lines under the formula "larger observations" should be
"bigger scales."

p. 22, 1st para., No. 3 last line should be simulation, not
stimulation.

p. 23. 1st para., line 7. "Table 9" should read "Table 6."

p. 44. Temperature table in second column on the page.

The temperatures read 13.9
 13.5
 16.2

The correct temperatures are 13.9
 15.2
 16.2.



- p.45. Second column, $Y_{i,t,k}$ should read $y_{i,t}$.
- p.46 1st Para., line 7, change the word "spawning" to "tows, \hat{T} ".
- p.49. Table 5d. Strike out the words "within or" in the second line of the heading.
- p.55. 9th line from the bottom, x_1 should be x_i .
25. p.56. First para. second column, sixth line, 26 should read
- p.63. Under "Preservation" $Na_2H_2PO_4$ should be Na_2HPO_4 .
- p.93. In table 1, atretic state e, change $>$ to $<$.
- p.97. In the! formula after the second para. change $<$ to $>$.
- p.98. In the formula in the first column change $-Zt$ to $-Zt_h$.